

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2001, 14:04:33 ; Search time 28.81 Seconds
(without alignments)

26.440 Million cell updates/sec

Title: US-09-692-401-5

Sequence: 1 VVRIGHLYIL 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

hed: 219241 segs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pIR1:*
2: pIR2:*
3: pIR3:*
4: pIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	314	2	melanoma antigen M
2	38	76.0	233	2	protein F23N19.5 (
3	37	74.0	314	2	melanoma antigen M
4	36	72.0	423	2	CDP-diacylglycerol
5	36	72.0	423	2	CDP-diacylglycerol
6	36	72.0	427	2	CDP-diacylglycerol
7	35	70.0	234	2	probable membrane
8	35	70.0	305	2	hypothetical prote
9	35	70.0	325	2	hypothetical prote
10	35	70.0	505	2	cytochrome P450 (E
11	35	70.0	613	2	GCEP protein (lipo
12	35	70.0	613	2	GCEP protein (lipo
13	35	70.0	637	2	hypothetical prote
14	35	70.0	728	1	hepatocyte growth
15	35	70.0	728	1	hepatocyte growth
16	35	70.0	728	1	hepatocyte growth
17	35	70.0	808	2	protein F2235.28 (
18	34	68.0	304	2	conserved hypotet
19	34	68.0	340	2	brain link protein
20	34	68.0	424	2	probable phosphat
21	34	68.0	438	2	CDP-diacylglycerol
22	34	68.0	1244	2	hypothetical prote
23	34	68.0	1288	2	hypothetical prote
24	34	68.0	1461	2	probable retrolem
25	33	66.0	101	2	conserved hypotet
26	33	66.0	168	2	hypothetical prote
27	33	66.0	313	2	calcium-binding pr
28	33	66.0	314	2	melanoma antigen M
29	33	66.0	337	2	probable ethylene-

30	33	66.0	538	1	COECD7	tuse protein - Esc
31	33	66.0	619	1	S48729	glucokinase regula
32	33	66.0	663	2	A75461	DNA gyrase, subun
33	33	66.0	1112	2	D75056	cell division cont
34	33	66.0	1232	2	D70615	multidrug resist
35	33	66.0	191	2	H84544	hypothetical prote
36	32	64.0	252	1	H69002	conserved hypotet
37	32	64.0	256	2	S26742	tonoplast intrinsi
38	32	64.0	256	2	J01106	hypothetical prote
39	32	64.0	321	2	F84611	N-acetylglucosamin
40	32	64.0	350	2	E72352	hypothetical prote
41	32	64.0	354	2	T04779	proximal sequence
42	32	64.0	368	2	JC6081	conserved hypotet
43	32	64.0	379	1	C70124	heterogeneous nucl
44	32	64.0	415	2	S43484	hypothetical prote
45	32	64.0	449	2	T29366	

ALIGNMENTS

RESULT 1

154519 melanoma antigen MAGE-12 - human

N:Alternate names: MAGE-21 protein

C:Species: Homo sapiens (man)

C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 18-Feb-2000

C:Accession: I54519; J02362; PH1295

R:De Smet, C.; Lurquin, C.; van der Bruggen, P.; De Plaen, E.; Brasseur, F.; Boon, T.

Immunogenetics 39, 121-129, 1994

A:Title: Sequence and expression pattern of the human MAGE-1 gene.

A:Reference number: I54519; MUID:94102805

A:Accession: I54519

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-314 <DES>

A:Cross-References: GB:I18877; NID:949345; PIDN:AAA19023.1; PID:949346

R:Braverman, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin, C.; Chomez, P.; Van P

J. Exp. Med. 176, 1453-1457, 1992

A:Title: A nonpeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytol

A:Reference number: PH1294; MUID:93018875

A:Accession: J02362

A:Molecule type: mRNA

A:Residues: 1-9, 'S', 'I', '186', 'D', '188-299', 'S', '301-314 <DIN>

A:Experimental source: melanoma cell line DM150; MAGE-12f

R:Braverman, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin, C.; Chomez, P.; Van P

J. Exp. Med. 176, 1453-1457, 1992

A:Title: A nonpeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytol

A:Reference number: PH1294; MUID:93018875

A:Accession: J02362

A:Molecule type: DNA

A:Residues: 1-9, 'S', 'I', '186', 'D', '188-299', 'S', '301-314 <DIN>

Query Match 100.0%; Score 50; DB 2; Length 314;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VVRIGHLYIL 10
DB 169 VVRIGHLYIL 178
RESULT 2
B96652
protein F23N19.5 [Imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: B96652
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Holt, D.; Blee, M.; Cooper, M.K.; Condon, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.; Jansen, R.K.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 C:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzall, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: B96652
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-233 <STO>
 C:Cross-references: GB:AE005173; NID:g6630448; PIDN:AA19536.1; GSPDB:GN00141
 C:Genetics: F23N19.5
 A:Gene: F23N19.5
 A:Map position: 1

Query Match 76.0%; Score 38; DB 2; Length 233;
 Best Local Similarity 75.0%; Pred. No. 3.7;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVRIGHLY 8
 :|||||:
 Db 176 IVRIGHY 183

RESULT 3
 168889
 melanoma antigen MAGE-2 - human
 C:Species: Homo sapiens (man)
 C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 18-Feb-2000
 C:Accession: I68889; PH1294
 R:De Smet, C.; Lurquin, C.; van der Bruggen, P.; De Plaen, E.; Brasseur, F.; Boon, T.
 Immunogenetics 39, 121-129, 1994
 A:Title: Sequence and expression pattern of the human MAGE2 gene.
 A:Reference number: I54519; MUID:94102805
 A:Accession: I68889
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-314 <RES>
 A:Cross-references: GB:U18920; NID:g436180; PIDN:AA17729.1; PID:g436181
 R:Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin, C.; Chomez, P.; Van Pel, J.; Exp. Med. 176, 1453-1457, 1992
 A:Title: A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytolytic T cells.
 A:Reference number: PH1294; MUID:93018875
 A:Accession: PH1294
 A:Molecule type: DNA
 A:Residues: 168-176 <TRA>
 C:Genetics:
 A:Gene: GDB:MAGE2; MAGE2
 A:Cross-references: GDB:273684
 A:Map position: XG28-XG28
 C:Superfamily: tumor associated protein MAGE

Query Match 74.0%; Score 37; DB 2; Length 314;
 Best Local Similarity 80.0%; Pred. No. 8.3;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 VVRIGHLY 10
 :|||||:
 Db 169 VVPISHLY 178

RESULT 4
 G85255

CDP-diacylglycerol synthetase-like protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
 C:Accession: G85255
 R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold SP
 Nature 402, 769-777, 1999
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A:Reference number: A85001; MUID:20083488
 A:Accession: G85255
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-423 <STO>
 A:Cross-references: GB:NC_001268; NID:g7269080; PIDN:CA879189.1; GSPDB:GN00140
 C:Genetics:
 A:Gene: AT4g22340
 A:Map position: 4

Query Match 72.0%; Score 36; DB 2; Length 423;
 Best Local Similarity 66.7%; Pred. No. 18;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VVRIGHLY 9
 :|||||:
 Db 66 IVYIGHLY 74

RESULT 5
 704915
 CDP-diacylglycerol synthetase homolog T10114.170 - Arabidopsis thaliana (fragment)
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
 C:Accession: T04915
 R:Beyan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.
 submitted to the Protein Sequence Database, April 1998
 A:Reference number: 215389
 A:Accession: T04915
 A:Molecule type: DNA
 A:Residues: 1-423 <BEV>
 A:Cross-references: EMBL:AL021712
 A:Experimental source: cultivar Columbia; BAC clone T10114
 C:Genetics:
 A:Map position: 4
 A:Insertions: 25/3; 110/2; 161/1; 208/2; 260/3; 276/3; 305/3; 348/3; 367/3; 395/3
 A:Note: T10114.170

Query Match 72.0%; Score 36; DB 2; Length 423;
 Best Local Similarity 66.7%; Pred. No. 18;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VVRIGHLY 9
 :|||||:
 Db 66 IVYIGHLY 74

RESULT 6
 T01455
 CDP-diacylglycerol synthetase homolog F2401.17 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 23-Mar-2001
 C:Accession: T01455
 R:Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.;
 et al.; Ecker, J.R.
 submitted to the EMBL Data Library, January 1998
 A:Description: Genomic sequence for Arabidopsis thaliana BAC F2401.
 A:Reference number: 214211
 A:Accession: T01455
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-427 <SHI>
 A:Cross-references: EMBL:AC003113; NID:g2689438; PID:g2781361
 A:Experimental source: cultivar Columbia

C:Genetics:
 A:Map position: 1
 A:Introns: 25/3; 110/2; 161/1; 208/2; 260/3; 276/3; 305/3; 348/3; 373/3; 401/3
 A:Note: P2401.17

Query Match 72.0%; Score 35; DB 2; Length 427;
 Best Local Similarity 77.8%; Pred. No. 18;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VVRIGHLYI 9
 ||:|||||
 Db 66 VVYMGHLYI 74

RESULT 7
 T37141
 Description: Membrane protein - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T37141
 R:Smith, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 Submitted to the EMBL Data Library, August 1999
 A:Reference number: Z21622
 A:Accession: T37141
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-234 <HAR>
 A:Cross-references: EMBL:AL109772; PIDN:CAB53266.1; GSPDB:GN00070; SCODEB:SCJ9A.05C
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCODEB:SCJ9A.05C

Query Match 70.0%; Score 35; DB 2; Length 234;
 Best Local Similarity 55.6%; Pred. No. 15;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VRIGHTLYI 10
 :||||:|
 Db 148 LEIGHMYVL 156

RESULT 8
 T40565
 Description: Hypothetical protein 4.60 - Bacillus subtilis plasmid pTA1060
 C:Species: Bacillus subtilis
 C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 24-Nov-1999
 C:Accession: T40565
 R:Meijer, W.J.; Venema, G.; Bron, S.
 Nucleic Acids Res. 23, 612-619, 1995
 A:Title: Characterization of single strand origins of cryptic rolling-circle plasmids from
 A:Reference number: T40549; MUID:95206941.
 A:Accession: T40565
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-305 <RES>
 A:Cross-references: EMBL:U32380; NID:g1049123; PIDN:AAC44422.1; PID:g1049128
 A:Experimental source: plasmid pTA1060
 C:Genetics:
 A:Superfamily: Bacillus subtilis plasmid pTA1060 hypothetical protein 4.60

Query Match 70.0%; Score 35; DB 2; Length 305;
 Best Local Similarity 66.7%; Pred. No. 20;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VRIGHTLYI 10
 ||||:|
 Db 197 VRIGHTLYI 205

RESULT 9
 T12852
 Description: Hypothetical protein yopr - Bacillus subtilis phage SPBC2

C:Species: Bacillus subtilis phage SPBC2
 C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 20-Jun-2000
 C:Accession: T12852; H69917
 R:Lazarovic, V.; Duesterhoeft, A.; Soldo, B.; Hilbert, H.; Maue, C.; Karamata, D.
 Submitted to the EMBL Data Library, August 1997
 A:Description: The complete nucleotide sequence of the Bacillus subtilis SPB2 phage
 A:Reference number: Z17583
 A:Accession: T12852

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-325 <LAZ>
 A:Cross-references: EMBL:AF020713; NID:g3025478; PID:g3025566; PIDN:AAC13061.1
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
 A.; Ehrlich, S.D.; Emmerich, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gal
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Huijo, M
 Koether, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
 A:Authors: Lauber, J.; Lazarovic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl
 A:Authors: Schlach, S.; Schroeder, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Se
 akeuchi, M.; Yamakoshi, A.; Yamada, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uehida
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchini, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
 A:Reference number: A69580; MUID:98044033
 A:Accession: H69917

A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-325 <KUN>
 A:Cross-references: GB:Z99114; GB:Z99115; GB:AL009126; NID:g2634478; PIDN:CAB13997.1;
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yopr
 C:Superfamily: Bacillus subtilis phage SPBC2 hypothetical protein yopr

Query Match 70.0%; Score 35; DB 2; Length 325;
 Best Local Similarity 75.0%; Pred. No. 22;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VVRIGHLYI 8
 :|||
 Db 256 IVRSGLYI 263

RESULT 10
 T10896
 Description: Cytochrome P450 (EC 1.14.-.-) 81B1c - Jerusalem artichoke
 C:Species: Helianthus tuberosus (Jerusalem artichoke)
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Feb-2001
 C:Accession: T10896
 R:Cabelli-Hurtado, F.; Batard, Y.; Salun, J.; Durst, F.; Pinot, F.; Werck-Reichhart,
 J. Biol. Chem. 273, 7260-7267, 1998
 A:Title: Cloning, expression in yeast and functional characterization of CYP81B1, a p
 A:Reference number: Z17204; MUID:98184826
 A:Accession: T10896
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-505 <CAB>
 A:Cross-references: EMBL:AJ000477; NID:g3059128; PIDN:CA04116.1; PID:g3059129
 A:Experimental source: cv. blanc commun
 C:Genetics:
 A:Gene: CYP81B1c
 C:Function:
 A:Description: Specifically catalyzes the hydroxylation of medium chain saturated fat
 C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
 C:Keywords: Chromoprotein; fatty acid metabolism; heme; iron; metalloprotein; oxidore

F:301.463/Domain: cytochrome P450 homology <P45>
F:441/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 70.0%; Score 35; DB 2; Length 505;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 IGHLYL 10
|||||:
Db 40 IGHLYL 46

RESULT 11

C86537
Gcpe protein [imported] - Chlamydia pneumoniae (strain J138)

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001

C:Accession: C86537

R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is

Nucleic Acids Res. 28, 2311-2314, 2000

A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

A:Reference number: A86491; MUID:20330349

A:Accession: C86537

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-613 <STO>

A:Cross-references: GB:BA000008; NID:98978745; PIDN:BA08581.1; GSPDB:GN00142

A:Experimental source: strain J138

C:Genetics:

A:Gene: gcpe

Query Match 70.0%; Score 35; DB 2; Length 613;
Best Local Similarity 87.5%; Pred. No. 44;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 VRIGHLYI 9
|||||:
Db 18 VRIGHLYI 25

RESULT 12

E72087
gcpe protein CP0383 [imported] - Chlamydia pneumoniae (strains CML029 and AK39)

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000

C:Accession: E72087; D81582

R:Kelman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.

Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; MUID:99206606

A:Accession: E72087

A:Molecule type: DNA

A:Residues: 1-613 <AN>

A:Cross-references: GB:AE001621; GB:AE001363; NID:94376641; PIDN:AAID8517.1; PID:9437665

A:Experimental source: strain CML029

R:Reed, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,

C.; Dodson, R.; Gwin, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MOPN and Chlamydia pneumoniae AR39.

A:Reference number: A81500; MUID:20150255

A:Accession: D81582

A:Molecule type: DNA

A:Residues: 1-613 <REA>

A:Cross-references: GB:AE002200; GB:AE002161; NID:97189305; PIDN:AAAF38230.1; PID:9718930

A:Experimental source: strain AK39, HL cells

C:Genetics:

A:Gene: gcpe; CP0383

Query Match 70.0%; Score 35; DB 2; Length 613;
Best Local Similarity 87.5%; Pred. No. 44;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 VRIGHLYI 9
|||||:
Db 18 VRIGHLYI 25

RESULT 13

H70535
hypothetical protein RV0669c - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C:Accession: H70535

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

Rajandram, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987

A:Accession: H70535

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-637 <COL>

A:Cross-references: GB:295972; GB:AL123456; NID:93261790; PIDN:CAB0388.1; PID:921432

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV0669c

C:Superfamily: Mycobacterium tuberculosis hypothetical protein RV0669c

Query Match 70.0%; Score 35; DB 2; Length 637;
Best Local Similarity 60.0%; Pred. No. 46;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VYRIGHLYI 10
:|||||:
Db 403 VYRIGHLYI 412

RESULT 14

JH0579
hepatocyte growth factor precursor [validated] - human

M:Alternate names: hepatopoietin A; scatter factor

C:Species: Homo sapiens (man)

C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 08-Dec-2000

C:Accession: JH0579; JH0333; A41140; B36677; A33512; A39006; PH0114; A37796;

R:Seiki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.

Gene 102, 213-219, 1991

A:Title: Organization of the human hepatocyte growth factor-encoding gene.

A:Reference number: JH0579; MUID:91340155

A:Accession: JH0579

A:Molecule type: DNA

A:Residues: 1-728 <SEK>

A:Cross-references: DBJF:D90318

A:Note: the authors translated the codon GAA for residue 662 as Gly

R:Seiki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.

submitted to JFID, March 1991

A:Description: Organization of the human hepatocyte growth factor-encoding gene.

A:Reference number: JH0333

A:Accession: JH0333

A:Molecule type: DNA

A:Residues: 1-481, 'RT', 484-728 <SE2>

R:Weidner, K.M.; Arikaki, N.; Hartmann, G.; Vandekerckhove, J.; Weingart, S.; Rieder,

Proc. Natl. Acad. Sci. U.S.A. 88, 7001-7005, 1991

A:Title: Evidence for the identity of human scatter factor and human hepatocyte growt

A:Reference number: A41140; MUID:91334393

A:Accession: A41140

A:Molecule type: mRNA

A:Residues: 1-728 <WEI>

A:Cross-references: GB:J73239; NID:9337935; PIDN:AAA64239.1; PID:9337936

R:Seiki, T.; Ihara, I.; Sugimura, A.; Shimonishi, M.; Nishizawa, T.; Asami, O.; Hagiya

Biochem. Biophys. Res. Commun. 172, 321-327, 1990

A>Title: Isolation and expression of cDNA for different forms of hepatocyte growth factor
 A/Reference number: A36677; MUID:91025062
 A/Accession: B36677
 A/Molecule type: mRNA
 A/Residues: 1-728 <SE3>
 A/Cross-references: GB:M60718; NID:q184031; PIDN:AAA52648.1; PID:q184032
 A/Accession: A36677
 A/Molecule type: mRNA
 A/Residues: 1-161,167-728 <SE4>
 A/Cross-references: EMBL:X16323
 A/Experimental source: leukocyte
 R/Miyazawa, K.; Tsudouchi, H.; Naka, D.; Takahashi, K.; Okigaki, M.; Arakaki, N.; Nakaya
 Blochem. Biophys. Res. Commun. 183, 967-973, 1989
 A>Title: Molecular cloning and sequence analysis of cDNA for human hepatocyte growth fac
 A/Reference number: A33512; MUID:89392017
 A/Accession: A33512
 A/Status: not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-728 <M1Y>
 A/Cross-references: GB:M29145; NID:q184041; PIDN:AAA52650.1; PID:q306846
 R/Rubin, J.S.; Chan, A.M.L.; Bottaro, D.P.; Burgess, W.H.; Taylor, W.G.; Cech, A.C.; Hilt
 Proc. Natl. Acad. Sci. U.S.A. 88, 415-419, 1991
 A>Title: A broad-spectrum human lung fibroblast-derived mitogen is a variant of hepatocy
 A/Reference number: A39006; MUID:91110540
 A/Accession: A39006
 A/Molecule type: mRNA
 A/Residues: 1-161,167-728 <RUB>
 A/Cross-references: GB:M55379
 A/Experimental source: embryonic lung
 R/Toshima, Y.; Arakaki, N.; Naka, D.; Takahashi, K.; Hirono, S.; Kondo, J.; Nakayama,
 Blochem. Biophys. Res. Commun. 175, 660-667, 1991
 A>Title: Identification of the N-terminal residue of the heavy chain of both native and
 A/Reference number: PH014; MUID:91207365
 A/Accession: PH014
 A/Molecule type: protein
 A/Residues: 32-43;53-58 <ROS>
 A/Experimental source: plasma
 R/Weidner, K.M.; Behrens, J.; Vandekerckhove, J.; Birchmeier, W.
 J. Cell Biol. 111, 2097-2108, 1990
 A>Title: Scatter factor: molecular characteristics and effect on the invasiveness of epi
 A/Reference number: A37796; MUID:91035621
 A/Accession: A37796
 A/Molecule type: protein
 A/Residues: 86-91;329-344;356-363,'XX',366-370;425-434;442-447,'X',449-450;543-546,'X',5
 R/Nakamura, T.; Nishizawa, T.; Hagiya, M.; Seki, T.; Shimonishi, M.; Sugimura, A.; Tashi
 Nature 342, 440-443, 1989
 A>Title: Molecular cloning and expression of human hepatocyte growth factor.
 A/Reference number: S06794; MUID:90066676
 A/Accession: S06794
 A/Molecule type: mRNA
 A/Residues: 1-31,'HK',34-77,'N',79-292,'V',294-299,'W',301-316,'A',318-335,'K',337-386,'
 A/Cross-references: EMBL:X16323; NID:q32081; PIDN:CAA34387.1; PID:q32082
 A/Experimental source: liver
 A/Note: The authors translated the codon CAG for residue 727 as Glu
 R/Hartman, G.; Naidani, L.; Weidner, K.M.; Sachs, M.; Vigna, E.; Comoglio, P.M.; Birchm
 Proc. Natl. Acad. Sci. U.S.A. 89, 11574-11578, 1992
 A>Title: A functional domain in the heavy chain of scatter factor/hepatocyte growth fact
 A/Reference number: 159214; MUID:93087571
 A/Accession: 159214
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-288,'ET' <HAR>
 A/Cross-references: GB:D02931; NID:q184033; PIDN:AAA52649.1; PID:q184034
 R/Miyazawa, K.; Kitamura, A.; Naka, D.; Kitamura, N.
 Eur. J. Biochem. 197, 15-22, 1991
 A>Title: An alternatively processed mRNA generated from human hepatocyte growth factor g
 A/Reference number: S15443; MUID:91200041
 A/Accession: S15443
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-288,'ET' <M1Y>
 A/Cross-references: EMBL:X57574; NID:q32083; PIDN:CAA40802.1; PID:q32084

R/Shima, N.; Nagao, M.; Ogaki, F.; Tsuda, E.; Murakami, A.; Higashio, K.
 Blochem. Biophys. Res. Commun. 180, 1151-1158, 1991
 A>Title: Tumor cytotoxic factor/hepatocyte growth factor from human fibroblasts: clon
 A/Reference number: 152253; MUID:92062058
 A/Accession: 152253
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 161-166 <SHI>
 A/Cross-references: GB:S62561; NID:q237996; PIDN:AA20169.1; PID:q237997
 A/Genetics:
 A:Gene: GDB:HGF
 A/Cross-references: GDB:127524; OMIM:142409
 A:Map position: 7q21.1-7q21.1
 A:Insertions: 30/1; 83/2; 161/2; 209/1; 249/2; 289/2; 347/2; 390/1; 424/2; 469/1;
 C/Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C/Function:
 A:Description: stimulates mitosis of hepatocytes and other cells
 A/Note: does not have proteinase activity
 C/Superfamily: hepatocyte growth factor; kringe homology; trypsin homology
 C/Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringe;
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-494/495-728/Product: hepatocyte growth factor #status experimental <ACH>
 F:32-494/Domain: alpha chain #status experimental <ACH>
 F:128-206/Domain: kringe homology <KR1>
 F:211-288/Domain: kringe homology <KR2>
 F:305-383/Domain: kringe homology <KR3>
 F:391-469/Domain: kringe homology <KR4>
 F:495-728/Domain: beta chain #status experimental <BCH>
 F:495-716/Domain: trypsin homology <TRY>
 F:32/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status expert
 F:194,402,566,653/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:487-604/Disulfide bonds: #status predicted

Query Match 70.0%; Score 35; DB 1; Length 728;
 Best Local Similarity 50.0%; Pred. No. 53;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VVRIGHLYTL 10
 Db 628 LIRVAHLIYM 637

RESULT 15
 A35644
 hepatocyte growth factor precursor - rat
 N/Alternate names: hepatietin A; scatter factor
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 21-Jul-2000
 C/Accession: A35644; S13211
 R/Tashiro, K.; Hagiya, M.; Nishizawa, T.; Seki, T.; Shimonishi, M.; Shimizu, S.; Naka
 Proc. Natl. Acad. Sci. U.S.A. 87, 3200-3204, 1990
 A>Title: Deduced primary structure of rat hepatocyte growth factor and expression of
 A/Reference number: A35644; MUID:90222197
 A/Accession: A35644
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-728 <TAS>
 A/Cross-references: GB:D90102; GB:M32987; NID:q220766; PIDN:BA14133.1; PID:q220767
 A/Note: The authors translated the codon GAG for residue 70 as Gln, GAC for residue 4
 R/Okajima, A.; Miyazawa, K.; Kitamura, N.
 Eur. J. Biochem. 193, 375-381, 1990
 A>Title: Primary structure of rat hepatocyte growth factor and induction of its mRNA
 A/Reference number: S13211; MUID:91031482
 A/Accession: S13211
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-728 <OKA>
 A/Cross-references: EMBL:X54400; NID:q56353; PIDN:CAA38266.1; PID:q4539554
 C/Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C/Function:
 A:Description: stimulates mitosis of hepatocytes and other cells
 A/Note: does not have proteinase activity

C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; pyridine
F:1-12/Domain: signal sequence #status predicted <SIG>
F:36-495/Product: hepatocyte growth factor #status predicted <GMT>
F:36-495/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
F:129-207/Domain: kringle homology <KR1>
F:212-289/Domain: kringle homology <KR2>
F:306-384/Domain: kringle homology <KR3>
F:392-470/Domain: kringle homology <KR4>
F:496-728/Domain: hepatocyte growth factor beta chain #status predicted <BOH>
F:496-719/Domain: trypsin homology <TRY>
F:33/Modified site: pyrrolidone carboxylate acid (Gln) (in mature form) #status predicted
F:295,403,559,656/Binding site: carboxylate (Asn) (covalent) #status predicted
F:488-607/Disulfide bonds: #status predicted

Query Match	70.0%;	Score 35;	DB 1;	Length 728;
Best Local Similarity	50.0%;	Pred. No. 53;		
Matches	5;	Conservative	4;	Mismatches 1; Indels 0; Gaps 0;

```
QY      1 VRIGHLYIL 10
          ::|: |||:
631 LRVVHLYIM 640
```

Search completed: June 20, 2001, 14:04:34
Job time: 81 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2001, 14:04:34 : Search time 28.81 Seconds
(without alignments)
21.152 Million cell updates/sec

Title: US-09-692-401-6
Perfect score: 42
Sequence: 1 RIGHTLYL 8

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	314	2	154519 melanoma antigen M
2	35	83.3	505	2	T10896 cytochrome P450 (E
3	34	81.0	234	2	T37141 probable membrane
4	34	81.0	1244	2	T01956 hypothetical prote
5	34	81.0	1461	2	E84389 probable retroel
6	33	78.6	305	2	I40565 melanoma antigen M
7	33	78.6	314	2	JC2361 glucokinase regula
8	33	78.6	619	1	S48729 hepatocyte growth
9	33	78.6	728	1	JH0579 hepatocyte growth
10	33	78.6	728	1	A35644 hepatocyte growth
11	33	76.2	330	2	E60185 N-acetylglucosamin
12	32	76.2	330	2	E72352 heterogenous nucl
13	32	76.2	415	2	S43484 heterogenous nucl
14	32	76.2	449	2	I39358 conserved hypotet
15	31	73.8	101	2	E82849 protein P23N19.5
16	31	73.8	233	2	B96652 hypothetical prote
17	31	73.8	276	2	T12552 melanoma antigen M
18	31	73.8	314	2	I68869 GDP-diacylglycerol
19	31	73.8	423	2	G85255 GDP-diacylglycerol
20	31	73.8	423	2	T04915 probable phosphom
21	31	73.8	465	2	C70594 UDP-N-acetylmuram
22	31	73.8	484	2	C84955 voltage-gated pota
23	31	73.8	497	2	JE0275 hypothetical prote
24	31	73.8	534	2	T27054 conserved hypotet
25	31	73.8	537	2	H84642 conserved hypotet
26	31	73.8	543	2	H81102 hypothetical prote
27	31	73.8	543	2	F81838 DNA topoisomerase
28	31	73.8	599	2	D70104 Gcpe protein limpo
29	31	73.8	613	2	C86537

30	31	73.8	613	2	E72087 gCpe protein CP036
31	31	73.8	615	2	C75278 DNA polymerase III
32	31	73.8	720	2	E82384 ornithine decarbox
33	31	73.8	1507	2	A40228 neurexin I-alpha p
34	31	73.8	1530	2	I45944 neurexin I-alpha -
35	30	71.4	65	2	A23783 hypothetical prote
36	30	71.4	144	2	G70023 hypothetical prote
37	30	71.4	184	2	T36644 probable transcrip
38	30	71.4	188	2	T38223 probable adenine p
39	30	71.4	189	2	A83573 conserved hypotet
40	30	71.4	247	2	S67685 hypothetical prote
41	30	71.4	256	2	S26742 tonoplast intrinsi
42	30	71.4	256	2	JQ1106 thiamin-phosphate
43	30	71.4	280	2	A75614 hypothetical prote
44	30	71.4	288	2	G64000 gene V protein - A
45	30	71.4	303	2	C32252

ALIGNMENTS

RESULT 1
154519 melanoma antigen MAGE-12 - human
N: Alternate names: MAGE 21 protein
C: Species: Homo sapiens (man)
C: Date: 07-Jun-1996 #sequence_revision (07-Jun-1996#text_change 18-Feb-2000
C: Accession: 154519, JC2362; PH1295
I: De Smet, C.; Lurquin, C.; van der Bruggen, P.; De Plaen, E.; Brasseur, F.; Boon, T.
Immunogenetics 39, 121-129, 1994
A: Title: Sequence and expression pattern of the human MAGE2 gene.
A: Reference number: 154519; MUID:94102805
A: Accession: 154519
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-314 <DES>
A: Cross-references: GB:118877; NID:9499345; PIDN:AA19023.1; PID:9499346
R: Ding, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.
Biochem. Biophys. Res. Commun. 202, 549-555, 1994
A: Title: Cloning and analysis of MAGE-1-related genes.
A: Reference number: JC2358; MUID:94311935
A: Accession: JC2362
A: Molecule type: mRNA
A: Residues: 1-9, 'S', 'I', '186, 'D', '188-299, 'S', '301-314 <DIN>
R: Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin, C.; Chomez, P.; Van P.
J. Exp. Med. 176, 1453-1457, 1992
A: Title: A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytoly
A: Reference number: PH1294; MUID:93018875
A: Accession: PH1295
A: Molecule type: DNA
A: Residues: 168-176 <TRN>
A: Experimental source: MAGE-21
C: Genetics:
A: Gene: GDB:MAGEA12; MAGE12; MAGE-12f
A: Cross-references: GDB:331129
C: Map position: Xq28-Xq28
C: Superfamily: tumor associated protein MAGE
F: 168-176/Region: HLA-A1 binding #status Predicted

Query Match 100.0%: Score 42; DB 2; Length 314;
Best local Similarity 100.0%: Pred. No. 0.39; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

QY 1 RIGHTLYL 8
DB 171 RIGHTLYL 178

RESULT 2
T10896 cytochrome P450 (EC 1.14.-.-) 81B1c - Jerusalem artichoke

C:Species: Helianthus tuberosus (Jerusalem artichoke)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Feb-2001
C:Accession: T10896
R:Cabelllo-Hurtado, F.; Batard, Y.; Salaun, J.; Durst, F.; Pinot, F.; Merck-Reichardt, D.
J. Biol. Chem. 273, 7260-7267, 1998
A:Title: Cloning, expression in yeast and functional characterization of CYP81B1, a plant
A:Reference number: 217204; MUID:98184826
A:Accession: T10896
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-505 <CAB>
A:Cross-references: EMBL:AJ000477; NID:g3059128; PIDN:CAA04116.1; PID:g3059129
A:Experimental source: cv. blanc commun
C:Genetics:
A:Gene: CYP81B1
C:Function:
A:Description: Specifically catalyzes the hydroxylation of medium chain saturated fatty
C:Superfamily: human cytochrome P450 CYP2B6; cytochrome P450 homology
C:Keywords: chromoprotein; fatty acid metabolism; heme; iron; metalloprotein; oxidoreduc
F:301-463/Domain: cytochrome P450 homology <P45>
F:441/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 83.3%; Score 35; DB 2; Length 505;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 IGHLYIL 8
| | | | | | | |
DB 40 IGHLYIL 46

RESULT 3
T37141
probable membrane protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T37141
R:Harris, D.; Bentley, S.D.; Parkhill, J.; Barrrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21622
A:Accession: T37141
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-234 <HAR>
A:Cross-references: EMBL:AL109972; PIDN:CAB53266.1; GSPDB:GN00070; SCODEB:SCJ9A.05c
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCODEB:SCJ9A.05c

Query Match 81.0%; Score 34; DB 2; Length 234;
Best Local Similarity 71.4%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 IGHLYIL 8
| | | | | | | |
DB 150 IGHLYIL 156

RESULT 4
T01956
hypothetical protein T2L5.9 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 24-Mar-1999
C:Accession: T01956
R:Geisel, C.; Smith, A.; Le, T.
submitted to the EMBL Data Library, October 1998
A:Description: The sequence of A. thaliana T2L5.
A:Reference number: Z14470
A:Accession: T01956
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-1244 <GEI>
A:Cross-references: EMBL:AF096371; NID:g3695386; PID:g3695393
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 4
A:Introns: 75/1; 599/1; 706/3; 798/3; 906/1
A:Note: T2L5.9

Query Match 81.0%; Score 34; DB 2; Length 1244;
Best Local Similarity 75.0%; Pred. No. 83;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RIGHLYIL 8
| | | | | | | |
DB 520 RIGHLYIL 527

RESULT 5
E84589
probable retroelement pol polyprotein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: E84589
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: E84589
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1461 <STO>
A:Cross-references: GB:AE002093; NID:g4586028; PIDN:AMD2646.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g20460
A:Map position: 2
C:Superfamily: retrovirus-related polypeptide

Query Match 81.0%; Score 34; DB 2; Length 1461;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RIGHLYIL 8
| | | | | | | |
DB 544 RIGHLYIL 551

RESULT 6
I40565
hypothetical protein 4.60 - Bacillus subtilis plasmid pTA1060
C:Species: Bacillus subtilis
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 24-Nov-1999
C:Accession: I40565
R:Meijer, W.J.; Venema, G.; Bron, S.
Nucleic Acids Res. 23, 612-619, 1995
A:Title: Characterization of single strand origins of cryptic rolling-circle plasmids
A:Reference number: I40549; MUID:95206941
A:Accession: I40565
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-305 <RES>
A:Cross-references: EMBL:U33380; NID:g1049123; PIDN:AAC44422.1; PID:g1049128
A:Experimental source: plasmid pTA1060
C:Genetics:
A:Genome: plasmid
C:Superfamily: Bacillus subtilis plasmid pTA1060 hypothetical protein 4.60

Query Match 78.6%; Score 33; DB 2; Length 305;
Best Local Similarity 71.4%; Pred. No. 28;


```

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 IGHLYL 8
Db 199 IGHVYL 205

RESULT 7
JC2361
melanoma antigen MAGE-3 - human
N:Alternate names: MAGE 3 protein
C:Species: Homo sapiens (man)
C>Date: 20-Feb-1995 #sequence, revision 20-Feb-1995 #text_change 18-Feb-2000
C:Accession: JC2361; PH1296; I38438
R: Ding, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.
Biochem. Biophys. Res. Commun. 202, 549-555, 1994
A:Title: Cloning and analysis of MAGE-1-related genes.
A:Reference number: JC2358; MUID:94311935
A:Accession: JC2361
A:Molecule type: mRNA
A:Residues: 1-314 <DIN>
A:Experimental source: melanoma cell line DM150
R: Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin, C.; Chomez, P.; Van Pel,
J. Exp. Med. 176, 1453-1457, 1992
A:Title: A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytolytic
A:Reference number: PH1294; MUID:93018875
A:Accession: PH1296
A:Molecule type: DNA
A:Residues: 168-176 <TRA>
R: Gaugler, B.; Van den Eynde, B.; van der Bruggen, P.; Romero, P.; Gaforio, J.J.; De Pla
J. Exp. Med. 179, 921-930, 1994
A:Title: Human gene MAGE-3 codes for an antigen recognized on a melanoma by autologous c
A:Reference number: I38438; MUID:94157413
A:Accession: I38438
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-314 <RES>
A:Cross-references: EMBL:U03735; NID:g468825; PIDN:AAA17446.1; PID:g468826
C:Genetics:
A:Gene: MAGE-3
C:Superfamily: tumor associated protein MAGE
P:168-176/Region: HLA-A1 binding #status predicted

Query Match 78.6%; Score 33; DB 2; Length 314;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 IGHLYL 7
Db 172 IGHLYL 177

RESULT 8
S48729
glucokinase regulator - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 10-Sep-1999 #sequence, revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S48729; S49339
R: Veiga-da-Cunha, M.; Detheux, M.; Matelet, N.; van Schaftingen, E.
Eur. J. Biochem. 225, 43-51, 1994
A:Title: Cloning and expression of a Xenopus liver cDNA encoding a fructose-phosphate-1r
A:Reference number: S48729; MUID:95010134
A:Accession: S48729
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-619 <VEI>
A:Cross-references: EMBL:X80901; NID:g556677; PIDN:CAA56863.1; PID:g556678
C:Superfamily: glucokinase regulator

Query Match 78.6%; Score 33; DB 1; Length 619;
Best Local Similarity 100.0%; Pred. No. 62;

```

```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 GHLTYL 8
Db 327 GHLTYL 332

RESULT 9
JH0579
hepatocyte growth factor precursor [validated] - human
N:Alternate names: hepatietin A; scatter factor
C:Species: Homo sapiens (man)
C>Date: 17-Aug-1992 #sequence, revision 17-Aug-1992 #text_change 08-Dec-2000
C:Accession: JH0579; JH0333; A41140; B36677; A33512; A39006; PH0114; A37796;
R: Seki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.
Gene 102, 213-219, 1991
A:Title: Organization of the human hepatocyte growth factor-encoding gene.
A:Reference number: JH0579; MUID:91340155
A:Accession: JH0579
A:Molecule type: DNA
A:Residues: 1-728 <SEK>
A:Cross-references: DDBJ:D90318
A:Note: the authors translated the codon GAA for residue 662 as Gly
R: Seki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.
submitted to JIPD, March 1991
A:Description: Organization of the human hepatocyte growth factor-encoding gene.
A:Reference number: JH0333
A:Accession: JH0333
A:Molecule type: DNA
A:Residues: 1-481, 'Rr', 484-728 <SE2>
R: Weidner, K.M.; Arakaki, N.; Hartmann, G.; Vandekerckhove, J.; Weingart, S.; Rieder,
Proc. Natl. Acad. Sci. U.S.A. 88, 7001-7005, 1991
A:Title: Evidence for the identity of human scatter factor and human hepatocyte growt
A:Reference number: A41140; MUID:91334333
A:Accession: A41140
A:Molecule type: mRNA
A:Residues: 1-728 <WEI>
A:Cross-references: GB:M73339; NID:g337935; PIDN:AAA64239.1; PID:g337936
R: Seki, T.; Ihara, I.; Sugimura, A.; Shimonishi, M.; Nishizawa, T.; Asami, O.; Hagiya
Biochem. Biophys. Res. Commun. 172, 321-327, 1990
A:Title: Molecular cloning and expression of cDNA for different forms of hepatocyte growth fa
A:Reference number: A36677; MUID:91025062
A:Accession: B36677
A:Molecule type: mRNA
A:Residues: 1-728 <SE3>
A:Cross-references: GB:M60718; NID:g184031; PIDN:AAA52648.1; PID:g184032
A:Accession: A36677
A:Molecule type: mRNA
A:Residues: 1-161, 167-728 <SEA>
A:Cross-references: EMBL:X16323
A:Experimental source: Leukocyte
R: Miyazawa, K.; Tsubouchi, H.; Naka, D.; Takahashi, K.; Okigaki, M.; Arakaki, N.; Nak
Biochem. Biophys. Res. Commun. 163, 967-973, 1989
A:Title: Molecular cloning and sequence analysis of cDNA for human hepatocyte growth
A:Reference number: A33512; MUID:89392017
A:Accession: A33512
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-728 <MTY>
A:Cross-references: GB:M29145; NID:g184041; PIDN:AAA52650.1; PID:g306846
R: Rubin, J.S.; Chan, A.M.L.; Bottaro, D.P.; Burgess, W.H.; Taylor, W.G.; Cech, A.C.;
Proc. Natl. Acad. Sci. U.S.A. 88, 415-419, 1991
A:Title: A broad-spectrum human lung fibroblast-derived mitogen is a variant of hepat
A:Reference number: A39006; MUID:91110540
A:Accession: A39006
A:Molecule type: mRNA
A:Residues: 1-161, 167-728 <ROB>
A:Cross-references: GB:M55379
A:Experimental source: embryonic lung
R: Yoshiyama, Y.; Arakaki, N.; Naka, D.; Takahashi, K.; Hirono, S.; Kondo, J.; Nakayam
Biochem. Biophys. Res. Commun. 175, 660-667, 1991
A:Title: Identification of the N-terminal residue of the heavy chain of both native a
A:Reference number: PH0114; MUID:91207365

```

A:Accession: PH0114
 A:Molecule type: protein
 A:Residues: 32-43;53-58 <XOS>
 A:Experimental source: plasma
 R:Weidner, K.M.; Behrens, J.; Vandekerckhove, J.; Birchmeier, W.
 J. Cell Biol. 111, 2097-2108, 1990
 A:Title: Scatter factor: molecular characteristics and effect on the invasiveness of epi
 A:Reference number: A37796; MUID:91035621
 A:Accession: A37796
 A:Molecule type: protein
 A:Residues: 86-91;329-344;356-363, 'XX', 366-370;425-434;442-447, 'X', 449-450;543-546, 'X', 5
 R:Nakamura, T.; Nishizawa, T.; Hagiya, M.; Seki, T.; Shimomishi, M.; Sugimura, A.; Tash
 Nature 342, 440-443, 1989
 A:Title: Molecular cloning and expression of human hepatocyte growth factor.
 A:Reference number: S06794; MUID:90066676
 A:Accession: S06794
 A:Molecule type: mRNA
 A:Residues: 1-31, 'HK', 34-77, 'N', 79-292, 'V', 294-299, 'W', 301-316, 'A', 318-335, 'K', 337-386, '
 A:Cross-references: EMBL:X16323; NID:932081; PIDN:CAA34387.1; PID:932082
 A:Experimental source: liver
 A:Note: the authors translated the codon CAG for residue 727 as Glu
 A:Note: part of this sequence, including the amino end of both the alpha and beta chains
 R:Hartmann, G.; Naldini, L.; Weidner, K.M.; Sachs, M.; Vigna, E.; Comoglio, P.M.; Birchm
 Proc. Natl. Acad. Sci. U.S.A. 89, 11574-11578, 1992
 A:Title: A functional domain in the heavy chain of scatter factor/hepatocyte growth fact
 A:Reference number: I59214; MUID:93087571
 A:Accession: I59214
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-288, 'ET', <HAR>
 A:Cross-references: GB:I02931; NID:9184033; PIDN:AA52649.1; PID:9184034
 R:Myazawa, K.; Kitamura, A.; Naka, D.; Kitamura, N.
 Eur. J. Biochem. 197, 15-22, 1991
 A:Title: An alternatively processed mRNA generated from human hepatocyte growth factor g
 A:Reference number: S15443; MUID:91200041
 A:Accession: S15443
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-288, 'ET', <MY2>
 A:Cross-references: EMBL:X57574; NID:932083; PIDN:CAA40802.1; PID:932084
 R:Shim, N.; Nagao, M.; Ogaki, F.; Tsuda, E.; Murakami, A.; Higashio, K.
 Biochem. Biophys. Res. Commun. 180, 1151-1158, 1991
 A:Title: Tumor cytotoxic factor/hepatocyte growth factor from human fibroblasts: cloning
 A:Reference number: I52253; MUID:92062058
 A:Accession: I52253
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 161-166 <SHI>
 A:Cross-references: GB:S62561; NID:9237996; PIDN:AAB20169.1; PID:9237997
 A:Genetics:
 A:Gene: GDB:HGF
 A:Cross-references: GDB:127524; OMIM:142409
 A:Map position: 7q21.1-7q21.1
 A:Introns: 30/1; 85/2; 123/1; 161/2; 209/1; 249/2; 289/1; 347/2; 390/1; 424/2; 469/1; 48
 C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C:Function:
 A:Description: stimulates mitosis of hepatocytes and other cells
 A:Note: does not have proteinase activity
 C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology; pyr
 C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; pyr
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-494/Domain: alpha chain #status experimental <ACH>
 F:128-206/Domain: kringle homology <KR1>
 F:211-288/Domain: kringle homology <KR2>
 F:305-383/Domain: kringle homology <KR3>
 F:391-469/Domain: kringle homology <KR4>
 F:495-728/Domain: beta chain #status experimental <BCH>
 F:495-716/Domain: trypsin homology <TRY>
 F:32/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
 F:294, 402, 566, 653/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:487-604/disulfide bonds: #status predicted

Query Match 78.6%; Score 33; DB 1; Length 728;
 Best Local Similarity 62.5%; Pred. NO. 74;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 RIGHLYIL 8
 Db 630 RVAHLIYM 637
 RESULT 10
 A35644
 hepatocyte growth factor precursor - rat
 N:Alternate names: hepatoinetin A; scatter factor
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 21-Jul-2000
 C:Accession: A35644; S13211
 R:Tashiro, K.; Hagiya, M.; Nishizawa, T.; Seki, T.; Shimomishi, M.; Shimizu, S.; Naka
 Proc. Natl. Acad. Sci. U.S.A. 87, 3200-3204, 1990
 A:Title: Deduced primary structure of rat hepatocyte growth factor and expression of
 A:Reference number: A35644; MUID:90222197
 A:Accession: A35644
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-728 <TRAS>
 A:Cross-references: GB:D90102; GB:M32987; NID:9220766; PIDN:BA41413.1; PID:9220767
 A:Note: the authors translated the codon CAG for residue 70 as Gln, GAC for residue 4
 R:Okajima, A.; Miyazawa, K.; Kitamura, N.
 Eur. J. Biochem. 193, 375-381, 1990
 A:Title: Primary structure of rat hepatocyte growth factor and induction of its mRNA
 A:Reference number: S13211; MUID:91031482
 A:Accession: S13211
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-728 <OKA>
 A:Cross-references: EMBL:X54400; NID:956353; PIDN:CAA38266.1; PID:94539554
 A:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C:Function:
 A:Description: stimulates mitosis of hepatocytes and other cells
 A:Note: does not have proteinase activity
 C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
 C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle;
 F:1-32/Domain: signal sequence #status predicted <SIG>
 F:56-495/Product: hepatocyte growth factor #status predicted <MAT>
 F:56-495/Product: hepatocyte growth factor alpha chain #status predicted <ACH>
 F:129-207/Domain: kringle homology <KR1>
 F:212-289/Domain: kringle homology <KR2>
 F:306-384/Domain: kringle homology <KR3>
 F:392-470/Domain: kringle homology <KR4>
 F:496-728/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
 F:496-719/Domain: trypsin homology <TRY>
 F:33/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
 F:295, 403, 569, 656/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:488-607/disulfide bonds: #status predicted
 Query Match 78.6%; Score 33; DB 1; Length 728;
 Best Local Similarity 62.5%; Pred. NO. 74;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 RIGHLYIL 8
 Db 633 RVAHLIYM 640
 RESULT 11
 A60185
 hepatocyte growth factor precursor - mouse
 N:Alternate names: hepatoinetin A; scatter factor
 C:Species: Mus musculus (house mouse)
 C>Date: 03-Mar-1993 #sequence_revision 26-May-1994 #text_change 16-Jun-2000
 C:Accession: J02117; PC2064; A60185; S43416; S45521; S17173; S10966; I48756; J00231
 R:Sasaki, M.; Nishio, M.; Sasaki, T.; Enami, J.

Biochem. Biophys. Res. Commun. 199, 772-779, 1994
 A:Title: Identification of mouse mammary fibroblast-derived mammary growth factor as hep
 A:Reference number: JC2117; MUID:94183257
 A:Accession: JC2117
 A:Molecule type: mRNA
 A:Residues: 1-728 <SAS2>
 A:Cross-references: GB:D10212; NID:q220435; PIDN:BA01064.1; PID:q220436
 A:Experimental source: fibroblast, COS-1 cell
 A>Note: submitted to JRPD, May 1993
 A:Accession: PC2064
 A:Molecule type: Protein
 A:Residues: 496-504 <SA2>
 R:Rosen, E.M.; Meromsky, L.; Setter, E.; Vinter, D.W.; Goldberg, I.D.
 Proc. Soc. Exp. Biol. Med. 195, 34-43, 1990
 A:Title: Purified scatter factor stimulates epithelial and vascular endothelial cell mig
 A:Reference number: A60185; MUID:90377927
 A:Accession: A60185
 A:Molecule type: Protein
 A:Residues: X, 184-188, 'KX', 191-192, 'X', 194, 'XX', 197, 357-364, 'XX', 367, 375-377, 'E', 379, 'R', Liu, Y.; Michalopoulos, G.K.; Zarnegar, R.
 Biochim. Biophys. Acta 1216, 299-303, 1993
 A:Title: Molecular cloning and characterization of cDNA encoding mouse hepatocyte growth
 A:Reference number: S43416; MUID:94060105
 A:Accession: S43416
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-728 <LTD>
 A:Cross-references: EMBL:X72307
 R:Liu, Y.
 submitted to the EMBL Data Library, May 1993
 A:Reference number: S45521
 A:Accession: S45521
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-563, 'H', 565-728 <LIZ>
 A:Cross-references: EMBL:X72307
 R:Coffey, A.; Fellows, J.; Young, S.; Pappin, D.; Rahman, D.
 Biochem. J. 278, 35-41, 1991
 A:Title: Purification and characterization of biologically active scatter factor from r
 A:Reference number: S17173; MUID:91354223
 A:Accession: S17173
 A:Molecule type: Protein
 A:Residues: 496-517, 'T', 519 <COF>
 R:Gherardi, E.; Stoker, M.
 Nature 346, 228, 1990
 A:Title: Hepatocytes and scatter factor.
 A:Reference number: S10966; MUID:90326152
 A:Accession: S10966
 A:Status: preliminary
 A:Molecule type: Protein
 A:Residues: 496-507, 'X', 509-512, 'L', 514-516, 'X', 518-519 <NAT>
 R:Plaschke-Schutter, A.; Behrens, J.; Gherardi, E.; Birchmeier, W.
 J. Biol. Chem. 270, 830-836, 1995
 A:Title: Characterization of the scatter factor/hepatocyte growth factor gene promoter.
 A:Reference number: 148758; MUID:9512532
 A:Accession: 148758
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-30 <RES>
 A:Cross-references: EMBL:X81630; NID:9673451; PIDN:CA57286.1; PID:9673452
 C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C:Function:
 A:Description: stimulates mitosis of hepatocytes and other cells
 A>Note: does not have proteinase activity
 C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
 C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; pyr
 F:1-33/Domain: signal sequence #status predicted <SIG>
 F:56-495/Domain: product; hepatocyte growth factor #status predicted <MAT>
 F:129-207/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
 F:212-289/Domain: kringle homology <KR1>
 F:306-384/Domain: kringle homology <KR2>
 F:392-470/Domain: kringle homology <KR3>

F:496-728/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
 F:496-719/Domain: trypsin homology <TRY>
 F:33/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
 F:295,403,569,656/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:488-607/Disulfide bonds: #status predicted
 Query Match 76.6%; Score 32; DB 1; Length 728;
 Best Local Similarity 62.5%; Pred. No. 74;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 RIGHLYIL 8
 Db 633 RVAHLYIM 640
 RESULT 12
 E72352
 N:acetylglucosaminyl-phosphatidylinositol biosynthesis-related protein - Thermotoga m
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C:Accession: E72352
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic
 Garret, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
 C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
 A:Reference number: A72200; MUID:99287316
 A:Accession: E72352
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-350 <ARN>
 A:Cross-references: GB:AE001737; GB:AE000512; NID:q4981144; PIDN:AAD5708.1; PID:q498
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TM0624
 C:Superfamily: probable hexosyltransferase ylxN
 Query Match 76.2%; Score 32; DB 2; Length 350;
 Best Local Similarity 75.0%; Pred. No. 53;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 RIGHLYIL 8
 Db 11 RIGHLLIM 18
 RESULT 13
 S43484
 heterogeneous nuclear ribonucleoprotein F - human
 N:Alternate names: hnRNP F; nucleolin-like protein mcs94-1
 C:Species: Homo sapiens (man)
 C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 05-Nov-1999
 C:Accession: S43484
 R:Matunis, M.J.; Xing, J.; Dreyfuss, G.
 Nucleic Acids Res. 22, 1059-1067, 1994
 A:Title: The hnRNP F protein: unique primary structure, nucleic acid-binding properti
 A:Reference number: S43484; MUID:94203790
 A:Accession: S43484
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-415 <MAT>
 A:Cross-references: GB:L28010; NID:q452047; PIDN:AA37584.1; PID:q452048
 A>Note: It is uncertain whether Met-1 or Met-2 is the initiator
 R:McDonald, H.; Smalls, D.; Jenkins, H.; Adams, K.; Simpson, N.E.; Goodfellow, P.J.
 Genomics 13, 314-348, 1992
 A:Title: Identification and characterization of a gene at D10S94 in the MEN2A region.
 A:Reference number: A42005; MUID:92307667
 A:Accession: A42005
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 2-415 <MCD>

C:Genetics:
 A:Gene: GDB:HNRP
 A:Cross-references: GDB:5429144; OMIM:601037
 A:Map position: 10q11.21-10q11.22

Query Match 76.2%; Score 32; DB 2; Length 415;
 Best Local Similarity 85.7%; Pred. No. 64;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RIGHLYI 7
 |||||
 DB 175 RIGHRYI 181

RESULT 14
 I39358
 heterologous nuclear ribonucleoprotein H - human
 N:Alternate names: 49K protein; hnRNP H
 C:Species: Homo sapiens (man)
 C>Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 05-Nov-1999
 C:Accession: I39358
 R:Honore, B.; Rasmussen, H.H.; Vorum, H.; Dejgaard, K.; Liu, X.; Gromov, P.; Madsen, P.;
 J. Biol. Chem. 270, 28780-28789, 1995
 A>Title: Heterologous nuclear ribonucleoproteins H, H', and F are members of a ubiquitin
 A:Reference number: A57806; MUID:96081943
 A:Accession: I39358
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-449 <RES>
 A:Cross-references: GB:I22009; NID:g347313; PIDN:AAA91346.1; PID:g347314
 C:Genetics:
 A:Gene: GDB:HNRP1; GDB:DBP49
 A:Cross-references: GDB:5428597; GDB:626033; OMIM:601035
 A:Map position: 5q35.3-5q35.3

Query Match 76.2%; Score 32; DB 2; Length 449;
 Best Local Similarity 85.7%; Pred. No. 70;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RIGHLYI 7
 |||||
 DB 175 RIGHRYI 181

RESULT 15
 E82849
 conserved hypothetical protein XP0095 [Imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
 C:Accession: E82849
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717
 A>Note: for a complete list of authors see reference number A59328 below
 A:Accession: E82849
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-101 <SIM>
 A:Cross-references: GB:AE003863; GB:AE003849; NID:g9104863; PIDN:AAF82908.1; GSPDB:GN001
 A:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carer, H
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fromm
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak

A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
 M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XP0095
 C:Superfamily: Methanococcus jannaschii conserved hypothetical protein M0652

Query Match 73.8%; Score 31; DB 2; Length 101;
 Best Local Similarity 75.0%; Pred. No. 21;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 RIGHLYI 8
 |||||
 DB 79 RIGHYI 86

Search completed: June 20, 2001, 14:04:35
 Job time: 82 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2001, 14:09:31 ; Search time 16.63 Seconds
(without alignments)
16.479 Million cell updates/sec

Title: US-09-692-401-6

Perfect score: 42

Sequence: 1 RIGHTLYL 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Swissprot_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	42	100.0	314 1 MAGC_HUMAN	P43365 homo sapien
2	33	78.6	314 1 MAGC_HUMAN	P43357 homo sapien
3	33	78.6	619 1 GCKR_XENLA	Q91754 xenopus lae
4	33	78.6	728 1 HGF_HUMAN	P14210 homo sapien
5	33	78.6	728 1 HGF_MOUSE	Q08048 mus musculu
6	33	78.6	728 1 HGF_RAT	P17945 rattus norv
7	32	76.2	415 1 ROF_HUMAN	P52597 homo sapien
8	32	76.2	449 1 ROH1_HUMAN	P31943 homo sapien
9	32	76.2	449 1 ROH2_HUMAN	P52795 homo sapien
10	32	76.2	510 1 C312_DROME	O97966 drosophila
11	31	73.8	314 1 MAGC_HUMAN	P43356 homo sapien
12	31	73.8	484 1 MURC_BUCAL	P57310 buchnera ap
13	31	73.8	599 1 PARE_BORBU	Q59189 borrelia bu
14	31	73.8	613 1 GCPE_CHLPN	Q928h0 chlamydia p
15	31	73.8	692 1 GYRB_BABBA	P44281 bartonella
16	30	71.4	256 1 TIPA_PAAVU	P23958 phaseolus v
17	30	71.4	288 1 Y077_HAEIN	P43935 haemophilus
18	30	71.4	303 1 PQOB_ACTICA	P07779 actinobact
19	30	71.4	314 1 MAGC_HUMAN	P43360 homo sapien
20	30	71.4	403 1 TGT_ARCFU	O28787 archaeoglob
21	30	71.4	421 1 CDS1_ARATH	O04928 a phosphati
22	30	71.4	424 1 CDS1_SOLTU	O04940 s phosphati
23	30	71.4	459 1 NEK3_HUMAN	P51956 homo sapien
24	30	71.4	510 1 C933_SOYBN	O81973 glycine max
25	30	71.4	683 1 EFGI_THEMA	O92154 thermotoga
26	30	71.4	883 1 RPOI_BPT7	P00574 bacterioph
27	30	71.4	906 1 RPOI_BPK11	P18147 bacterioph
28	30	71.4	3898 1 POLG_BVDVS	Q01499 bovine vita
29	30	71.4	3988 1 POLG_BVDVS	Q01499 bovine vita
30	29	69.0	76 1 RPOA_ARCFU	P19711 bovine vita
31	29	69.0	167 1 LSPA_CHLTF	O29135 archaeoglob
32	29	69.0	230 1 VME1_CVHOC	O84413 chlamydia t
33	29	69.0	253 1 YTBQ_BACSV	Q01455 human coron
				P53560 bacillus su

34	29	69.0	275 1 AMOA_NITEU	Q04507 nitrosomona
35	29	69.0	290 1 EXOS_BPT5	P06225 bacterioph
36	29	69.0	321 1 FLIM_AGRTU	O44457 agrobacteri
37	29	69.0	368 1 SNCI_HUMAN	Q16533 homo sapien
38	29	69.0	370 1 KCCI_HUMAN	Q14012 homo sapien
39	29	69.0	374 1 KCCI_RAT	O63450 rattus norv
40	29	69.0	397 1 GCST_BOVIN	P25285 bos taurus
41	29	69.0	403 1 GCST_HUMAN	P48728 homo sapien
42	29	69.0	404 1 SGAH_HYPM	O08374 hyphomicrob
43	29	69.0	412 1 YG08_SYNY3	P73001 synchocyst
44	29	69.0	512 1 DNB2_ADE04	P06500 human adeno
45	29	69.0	542 1 XP55_STRLI	P06109 streptomyce

ALIGNMENTS

RESULT	1	ALIGNMENTS
AC	MAGC_HUMAN	STANDARD; PRT; 314 AA.
AC	P43365;	
DT	01-NOV-1995 (Rel. 32, Created)	
DT	01-NOV-1995 (Rel. 32, Last sequence update)	
DT	15-JUL-1999 (Rel. 38, Last annotation update)	
DE	MELANOMA-ASSOCIATED ANTIGEN 12 (MAGE-12 ANTIGEN) (MAGE12P).	
GN	MAGEA12 OR MAGE12.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
OX	NCBI_TaxID:9606;	
RP	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE:94102805; PubMed:8276455;	
RA	de Smet C., Lurquin C., van der Bruggen P., de Plaen E., Brasseur F.,	
RA	Boon T.;	
RT	*Sequence and expression pattern of the human MAGE2 gene.*;	
RL	Immunogenetics 39:121-129(1994).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE-SKIN;	
RX	MEDLINE:94311935; PubMed:8037761;	
RA	Ding M., Beck R.J., Keller C.J., Fenton R.G.;	
RT	*Cloning and analysis of MAGE-1-related genes.*;	
RL	Biochem. Biophys. Res. Commun. 202:549-555(1994).	
CC	-1- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN TUMOR	
CC	TRANSFORMATION OR ASPECTS OF TUMOR PROGRESSION.	
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,	
CC	SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG	
CC	CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT	
CC	FOR TESTES.	
CC	-1- SIMILARITY: BELONGS TO THE MAGE FAMILY. STRONG SIMILARITY (90%)	
CC	WITH MAGE-2.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement. (See http://www.isb-sdb.ch/announce/	
CC	or send an email to license@isb-sdb.ch).	
CC	-----	
DR	EMBL: L18877; AAA19023.1; -	
DR	MIT: 300177; -	
DR	InterPro: IPR002190; -	
DR	Pfam: PF01454; MAGE; 1.	
KW	Antigen; Multigene family; Tumor antigen.	
FT	DOMAIN 40 43 POLY-SER.	
FT	CONFICT 10 10 C -> S (IN REF. 2).	
FT	CONFICT 187 187 A -> D (IN REF. 2).	
FT	CONFICT 300 300 P -> S (IN REF. 2).	
SEQUENCE	314 AA: 34802 MW: 3f0787CECB8816A5 CRC64;	

Query Match 100.0%; Score 42; DB 1; Length 314;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIGHTYL 8
 DB 171 RIGHTYL 178

RESULT 2
 ID MAG3_HUMAN STANDARD; PRT; 314 AA.
 AC P43357;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE MELANOMA-ASSOCIATED ANTIGEN 3 (MAGE-3 ANTIGEN) (ANTIGEN MZ2-D).
 GN MAGE3 OR MAGE3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. AND MUTAGENESIS.
 RC TISSUE=Blood;
 RA MEDLINE=94157413; PubMed=8113684;
 RA Gaugler B., van den Eynde B., van der Bruggen P., Romero P.,
 RA Gattorio J.J., de Plaen E., Lethe B., Brasseur F., Boon T.,
 RT "Human gene MAGE-3 codes for an antigen recognized on a melanoma by
 RT autologous cytolytic T lymphocytes.";
 RL J. Exp. Med. 179:921-930(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RA MEDLINE=94311935; PubMed=8037761;
 RA Ding M., Beck R.J., Keller C.J., Fenton R.G.,
 RT "Cloning and analysis of MAGE-1-related genes.";
 RL Biochem. Biophys. Res. Commun. 202:549-555(1994).
 CC -1- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
 CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
 CC PROGRESSION. ANTIGEN RECOGNIZED ON A MELANOMA BY AUTOLOGOUS
 CC CYTOLYTIC T LYMPHOCYTES.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
 CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG,
 CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
 CC FOR TESTES AND PLACENTA. NEVER EXPRESSED IN KIDNEY TUMORS,
 CC LEUKEMIAS AND LYMPHOMAS.
 CC -1- SIMILARITY: BELONGS TO THE MAGE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U03735; AAA17446.1; -
 DR MIM: 300174;
 DR InterPro: IPR002190; -
 DR Pfam: PF01454; MAGE; 1.
 KW Antigen; Multigene family; Tumor antigen.
 FT DOMAIN 40 43 POLY-SER.
 FT MUTAGEN 170 170 D->A: ABOLISHES HLA-A1 BINDING.
 FT MUTAGEN 176 176 Y->A: ABOLISHES HLA-A1 BINDING.
 SQ SEQUENCE 314 AA; 34747 MW; 3F5EB13D1C9946A1 CRC64;

Query Match 78.6%; Score 33; DB 1; Length 314;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 IGHLYT 7

DB 172 IGHLYT 177

RESULT 3
 ID GCRK_XENLA STANDARD; PRT; 619 AA.
 AC 091754;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GLUCOKINASE REGULATORY PROTEIN (GLUCOKINASE REGULATOR).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA MEDLINE=95010134; PubMed=7925465;
 RA Velga-Da-Cunha M., Dethieux M., Watelet N., van Schaftingen E.,
 RT "Cloning and expression of a Xenopus liver cDNA encoding a fructose-
 RT phosphate-insensitive regulatory protein of glucokinase.";
 RL Eur. J. Biochem. 225:43-51(1994).
 CC -1- FUNCTION: INHIBITS GLUCOKINASE BY FORMING AN INACTIVE COMPLEX WITH
 CC THIS ENZYME.
 CC -1- SIMILARITY: BELONGS TO THE SIS FAMILY. GCRK SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X80901; CAA56863.1; -
 DR InterPro: IPR001347; -
 DR InterPro: IPR001741; -
 DR Pfam: PF01380; SIS; 1.
 DR PROSITE: PS01272; GCRK; 1.
 SQ SEQUENCE 619 AA; 68738 MW; 41B72C1981D1BA52 CRC64;

Query Match 78.6%; Score 33; DB 1; Length 619;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GHLYT 8
 DB 327 GHLYT 332

RESULT 4
 ID HGF_HUMAN STANDARD; PRT; 728 AA.
 AC P14210;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HEPATOCTYTE GROWTH FACTOR PRECURSOR (SCATTER FACTOR) (SF)
 DE (HEPATOPOIETIN A).
 GN HGF OR HPTA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=91340155; PubMed=1831432;
 RA Seki T., Hagiya M., Shimomishi M., Nakamura T., Shimizu S.,
 RT "Organization of the human hepatocyte growth factor-encoding gene.";

RL Gene 102:213-219(1991).
 [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Placenta;
 RX MEDLINE=89392017; PubMed=2528952;
 RA Miyazawa K., Tsubouchi H., Naka D., Takahashi K., Okigaki M.,
 Arakaki N., Nakayama H., Hiroo S., Sakiyama O., Takahashi K.,
 Goida E., Daikuhara Y., Kitamura N.;
 RT "Molecular cloning and sequence analysis of cDNA for human hepatocyte
 growth factor.";
 RL Biochem. Biophys. Res. Commun. 163:967-973(1989).
 [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-leukocyte;
 RX MEDLINE=91025062; PubMed=2145836;
 RA Seki T., Ihara I., Sugimura A., Shimonishi M., Nishizawa T.,
 Asami O., Hagiya M., Nakamura T., Shimizu S.;
 RT "Isolation and expression of cDNA for different forms of hepatocyte
 growth factor from human leukocyte.";
 RL Biochem. Biophys. Res. Commun. 172:321-327(1990).
 [4]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 55-73 AND 495-520.
 RC TISSUE-Liver;
 RX MEDLINE=9006676; PubMed=2531289;
 RA Nakamura T., Nishizawa T., Hagiya M., Seki T., Shimonishi M.,
 Sugimura A., Tashiro K., Shimizu S.;
 RT "Molecular cloning and expression of human hepatocyte growth factor.";
 RL Nature 342:440-443(1989).
 [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Embryonic fibroblast;
 RX MEDLINE=91334393; PubMed=1831266;
 RA Weidner K.M., Arakaki N., Hartman G., Vandekerckhove J., Weingart S.,
 Rieder H., Fornsusch C., Tsubouchi H., Hishida T., Daikuhara Y.,
 Birchmeier W.;
 RT "Evidence for the identity of human scatter factor and human
 hepatocyte growth factor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:7001-7005(1991).
 [6]
 RP SIGNAL SEQUENCE CLEAVAGE SITE
 RX MEDLINE=91207365; PubMed=1826837;
 RA Yoshiyama Y., Arakaki N., Naka D., Takahashi K., Hiroo S., Kondo J.,
 Nakayama H., Goida E., Kitamura N., Tsubouchi H., Ishii T.,
 Hishida T., Daikuhara Y.;
 RT "Identification of the N-terminal residue of the heavy chain of both
 native and recombinant human hepatocyte growth factor.";
 RL Biochem. Biophys. Res. Commun. 175:660-667(1991).
 [7]
 RP CARBOHYDRATE-LINKAGE SITE 476.
 RX MEDLINE=93129192; PubMed=1482348;
 RA Shimizu N., Hara H., Sogabe T., Sakai H., Ihara I., Inoue H.,
 Nakamura T., Shimizu S.;
 RT "Hepatocyte growth factor is linked by O-glycosylated oligosaccharide
 on the alpha chain.";
 RL Biochem. Biophys. Res. Commun. 189:1329-1335(1992).
 [8]
 RP MUTAGENESIS.
 RX MEDLINE=92331602; PubMed=1321034;
 RA Lokker N.A., Mark M.R., Luis E.A., Bennett G.L., Robbins K.A.,
 Baker J.B., Godowski P.J.;
 RT "Structure-function analysis of hepatocyte growth factor:
 identification of variants that lack mitogenic activity yet retain
 high affinity receptor binding.";
 RL EMBO J. 11:2503-2510(1992).
 [9]
 RP STRUCTURE BY NMR OF 31-127.
 RX MEDLINE=96154323; PubMed=9493272;
 RA Zhou H., Mazzulla M.J., Kaufman J.D., Stahl S.J., Wingfield P.T.,
 Rubin J.S., Bottaro D.P., Byrd R.A.;
 RT "The solution structure of the N-terminal domain of hepatocyte growth
 factor reveals a potential heparin-binding site.";
 RL Structure 6:109-116(1998).
 [10]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 35-210.
 RX MEDLINE=99036858; PubMed=9817840;
 RA Ullrich M., Lokker N.A., Godowski P.J., de Vos A.M.;
 RT "Crystal structure of the NK1 fragment of human hepatocyte growth
 factor at 2.0-A resolution.";
 RL Structure 6:1383-1393(1998).
 CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
 HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPIC FACTOR, AND ACTS
 AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
 CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.
 CC -1- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
 DISULFIDE BOND.
 CC -1- SIMILARITY: CONTAINS 4 KRINGLE REGIONS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 or send an email to license@isb-sdb.ch).
 CC -----
 DR EMBL; D90334; BAA14348.1; -;
 DR EMBL; D90318; BAA14348.1; JOINED.
 DR EMBL; D90319; BAA14348.1; JOINED.
 DR EMBL; D90320; BAA14348.1; JOINED.
 DR EMBL; D90322; BAA14348.1; JOINED.
 DR EMBL; D90323; BAA14348.1; JOINED.
 DR EMBL; D90324; BAA14348.1; JOINED.
 DR EMBL; D90325; BAA14348.1; JOINED.
 DR EMBL; D90326; BAA14348.1; JOINED.
 DR EMBL; D90327; BAA14348.1; JOINED.
 DR EMBL; D90328; BAA14348.1; JOINED.
 DR EMBL; D90329; BAA14348.1; JOINED.
 DR EMBL; D90330; BAA14348.1; JOINED.
 DR EMBL; D90331; BAA14348.1; JOINED.
 DR EMBL; D90332; BAA14348.1; JOINED.
 DR EMBL; D90333; BAA14348.1; JOINED.
 DR EMBL; M29143; AA52655.1; -;
 DR EMBL; M60718; AA52648.1; -;
 DR EMBL; X16323; CAA33387.1; -;
 DR EMBL; M73239; AAA64239.1; -;
 DR EMBL; M73240; AAA64297.1; -;
 DR PIR; JH0579; JH0579.
 DR PIR; S06794; S06794.
 DR PDB; 2HGF; 24-JUN-98.
 DR PDB; 1BHT; 18-NOV-98.
 DR MEROPS; S01.976; -;
 DR GLYCOSULEDB; P14210; -;
 DR MIM; 142409; -;
 DR InterPro; IPR000001; -;
 DR InterPro; IPR001254; -;
 DR InterPro; IPR001314; -;
 DR InterPro; IPR003014; -;
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00051; kringle; 4.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00018; KRINGLE.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PROSITE; PS00021; KRINGLE_1; 4.
 DR PROSITE; PS50070; KRINGLE_2; 4.
 KW Growth factor; Kringle; Glycoprotein; Serine protease homolog;
 KW Signal; 3D-structure
 FT SIGNAL 1 31
 FT CHAIN 32 494
 FT CHAIN 495 728
 FT MOD_RES 32 32
 FT DOMAIN 32 127
 FT DOMAIN 128 206
 FT DOMAIN 211 288
 FT DOMAIN 305 383
 FT
 HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.
 HEPATOCYTE GROWTH FACTOR BETA CHAIN.
 PYROGLUTAMINE CARBOXYLIC ACID.
 PAB.
 KRINGLE 1.
 KRINGLE 2.
 KRINGLE 3.

```

FT DOMAIN 391 469 KRINGLE 4.
FT DOMAIN 495 728 SERINE PROTEASE-LIKE.
FT DISULFID 70 96
FT DISULFID 74 84
FT DISULFID 128 206
FT DISULFID 149 189
FT DISULFID 177 201
FT DISULFID 487 604
FT CARBOHYD 294 294
FT CARBOHYD 402 402
FT CARBOHYD 476 476
FT CARBOHYD 566 566
FT CARBOHYD 653 653
FT CARBOHYD 728 728
FT CARBOHYD 78 78
FT CARBOHYD 293 293
FT CARBOHYD 300 300
FT CARBOHYD 317 317
FT CARBOHYD 336 336
FT CARBOHYD 387 387
FT CARBOHYD 416 416
FT CARBOHYD 505 505
FT CARBOHYD 509 509
FT CARBOHYD 558 558
FT CARBOHYD 561 561
FT CARBOHYD 595 595
SQ SEQUENCE 728 AA; 83133 MW; 2D997938295AD2F CRC64;

Query Match 78.6%; Score 33; DB 1; Length 728;
Best Local Similarity 62.5%; Pred. No. 34;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 RIGHTLYL 8
DB 630 RVAHLITM 637

RESULT 5
ID HGF_MOUSE STANDARD; PRT; 728 AA.
AC Q08048; Q64007; Q61662;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE HEPATOCYTE GROWTH FACTOR PRECURSOR (SCATTER FACTOR) (SF)
DE (HEPATOPOETIN A).
GN HGF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 496-504.
RC TISSUE=Mammary fibroblast;
RX MEDLINE=94183257; PubMed=8135822;
RA Sasaki M., Nishio M., Sasaki T., Enami J.;
RT "Identification of mouse mammary fibroblast-derived mammary growth
factor as hepatocyte growth factor."
RL Biochem. Biophys. Res. Commun. 199;772-779(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94363381; PubMed=8081873;
RA Lee C.C., Kozak C.A., Yamada K.M.;
RT "Structure, genetic mapping, and expression of the mouse Hgf/scatter
factor gene."
RL Cell Adhes. Commun. 1;101-111(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94060105; PubMed=8241272;
RA Liu Y., Michalopoulos G.K., Zarnegar R.;

```

```

RT "Molecular cloning and characterization of cDNA encoding mouse
hepatocyte growth factor."
RL Biochim. Biophys. Acta 1216:299-303(1993).
CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPIC FACTOR, AND ACTS
AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.
CC -1- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
DISULFIDE BOND.
CC -1- ALTERNATIVE PRODUCTS: A SHORT FORM OF HGF IS PRODUCED BY
ALTERNATIVE RNA SPLICING. THE SEQUENCE SHOWN HERE IS THAT OF THE
LONG FORM.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10212; BAA01064.1; -
DR EMBL; D10213; BAA01065.1; -
DR EMBL; S71816; AAB31855.1; -
DR EMBL; X72307; CAA51054.1; ALT_INIT.
DR HSSP; P14210; ZHGF.
DR MGD; MGI:96079; Hgf.
DR InterPro; IPR000001; -
DR InterPro; IPR001254; -
DR InterPro; IPR001314; -
DR InterPro; IPR003014; -
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS0070; KRINGLE_2; 4.
KW Growth factor; Kringle; Glycoprotein; Serine protease homolog;
KW Signal; Alternative splicing.
FT SIGNAL 1 32
FT CHAIN 33 495
FT CHAIN 496 728
FT MOD_RES 33 33
FT DOMAIN 33 128
FT DOMAIN 129 207
FT DOMAIN 212 289
FT DOMAIN 305 384
FT DOMAIN 392 470
FT DOMAIN 496 728
FT DISULFID 71 97
FT DISULFID 75 85
FT DISULFID 163 163
FT DISULFID 488 607
FT CARBOHYD 295 295
FT CARBOHYD 403 403
FT CARBOHYD 469 469
FT CARBOHYD 566 566
FT CARBOHYD 656 656
FT CARBOHYD 728 728
FT VARSPIC 163 167
FT CONFLICT 344 344
FT CONFLICT 479 479
FT CONFLICT 564 564
SQ SEQUENCE 728 AA; 82944 MW; A0381FC497534528 CRC64;

Query Match 78.6%; Score 33; DB 1; Length 728;
Best Local Similarity 62.5%; Pred. No. 34;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 RIGHTLYL 8

```



```

DB      633 RVALHYIM 640

RESULT 6
HGF_RAT
ID HGF_RAT STANDARD: PRT: 728 AA.
AC P17945;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HEPATOCYTE GROWTH FACTOR PRECURSOR (SCATTER FACTOR) (SF)
   (HEPATOPOEITIN A).
GN HGF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OK NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=WISTAR; TISSUE=Liver;
RX MEDLINE=90222197; PubMed=2139229;
RA Toshiro K., Hagiya M., Nishizawa T., Seki T., Shimonishi M.,
RA Shimitzu S., Nakamura T.;
RT "Deduced primary structure of rat hepatocyte growth factor and
RT expression of the mRNA in rat tissues."
RL Proc. Natl. Acad. Sci. U.S.A. 87:3200-3204(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR; TISSUE=Liver;
RX MEDLINE=91031482; PubMed=2146117;
RA Okajima A., Miyazawa K., Kitamura N.;
RT "Primary structure of rat hepatocyte growth factor and induction of
RT its mRNA during liver regeneration following hepatic injury."
RL Eur. J. Biochem. 193:375-381(1990).
CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
CC HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPHIC FACTOR, AND ACTS
CC AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.
CC -1- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
CC DISULFIDE BOND.
CC -1- SIMILARITY: CONTAINS 4 KRINGLE REGIONS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY. PLASMINOGEN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D90102; BAA14133.1; -
DR EMBL: X54400; CA38266.1; -
DR PIR: S13211; S13211.
DR PIR: A35644; A35644.
DR HSSP: P14210; 2HGF.
DR MEROPS: S01.978; -
DR InterPro: IPR000001; -
DR InterPro: IPR001254; -
DR InterPro: IPR001314; -
DR InterPro: IPR003014; -
DR Pfam: PF00024; PAN.1.
DR Pfam: PF00051; kringle.4.
DR Pfam: PF00089; trypsin.1.
DR PRINTS: PR00018; KRINGLE.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PROSITE: PS00021; KRINGLE_1; 4.
DR PROSITE: PS00700; KRINGLE_2; 4.
RW Growth factor; Kringle; Glycoprotein; Serine protease homolog;
FT SIGNAL 1 32 BY SIMILARITY.

```

```

FT CHAIN 33 495 HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.
FT CHAIN 496 728 HEPATOCYTE GROWTH FACTOR BETA CHAIN.
FT MOD_RES 33 33 PYROGLUTAMINE CARBOXYLIC ACID
   (BY SIMILARITY).
FT DOMAIN 33 128 PAP.
FT DOMAIN 129 207 KRINGLE 1.
FT DOMAIN 212 289 KRINGLE 2.
FT DOMAIN 306 384 KRINGLE 3.
FT DOMAIN 392 470 KRINGLE 4.
FT DOMAIN 496 728 SERINE PROTEASE-LIKE.
FT DISULFID 71 97 BY SIMILARITY.
FT DISULFID 75 85 BY SIMILARITY.
FT DISULFID 488 607 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 656 656 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 728 AA; 82905 MW; 3E0BF1F96ADCDEF CRC64;

Query Match
Best Local Similarity 78.6%; Score 33; DB 1; Length 728;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RIGHLYIL 8
Db 633 RVALHYIM 640

RESULT 7
ROF_HUMAN
ID ROF_HUMAN STANDARD: PRT: 415 AA.
AC P52597;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN F (HNRNP F).
GN HNRPF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OK NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 51-60; 77-82 AND 127-135.
RX MEDLINE=94203790; PubMed=7512260;
RA Matunis M.J., Xing J., Dreyfuss G.;
RT "The hnrp f protein: unique primary structure, nucleic acid-binding
RT properties, and subcellular localization."
RL Nucleic Acids Res. 22:1059-1067(1994).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96081943; PubMed=7499401;
RA Honore B., Rasmussen H.H., Vorum H., Dejgaard K., Liu X.,
RA Gromov P., Madsen P., Gesser B., Tommerup N., Celis J.E.;
RT "Heterogeneous nuclear ribonucleoproteins H, H', and F are members of
RT a ubiquitously expressed subfamily of related but distinct proteins
RT encoded by genes mapping to different chromosomes."
RL J. Biol. Chem. 270:28780-28789(1995).
CC -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE HETEROGENEOUS NUCLEAR
CC RIBONUCLEOPROTEIN (HNRNP) COMPLEXES WHICH PROVIDE THE SUBSTRATE
CC FOR THE PROCESSING EVENTS THAT PRE-MRNAs UNDERGO BEFORE BECOMING
CC FUNCTIONAL, TRANSLATABLE MRNAs IN THE CYTOPLASM. PROBABLY BINDS G-
CC RICH SEQUENCES IN PRE-MRNAs.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC -1- TISSUE SPECIFICITY: EXPRESSED UBICITOUSLY.
CC -1- SIMILARITY: CONTAINS 3 RNA RECOGNITION MOTIFS (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@lsb.sib.ch).
CC -----
DR EMBL: L28010; AAC37584.1; -.
DR Aarhus/Chent-2DPAGE; 6304; IEF.
DR Aarhus/Chent-2DPAGE; 7312; IEF.
DR MIM: 601037; -.
DR InterPro: IPR000504; -.
DR Pfam: PF00076; rrm; 3.
DR PROSITE: PS50102; RRM; 2.
DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat.
FT DOMAIN 11 188 RNA-BINDING (RRM) 1.
FT DOMAIN 289 366 RNA-BINDING (RRM) 2.
FT DOMAIN 366 RNA-BINDING (RRM) 3.
SQ SEQUENCE 415 AA; 45672 MW; D14E170631FB1F31 CRC64;

Query Match
Best Local Similarity 85.7%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIGHLYI 7
Db 175 RIGHRYI 181

RESULT 8
ID ROH1_HUMAN STANDARD; PRT; 449 AA.
AC P31943;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1996 (Rel. 34, last sequence update)
DT 01-OCT-2000 (Rel. 40, last annotation update)
DE HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN H (HNRNP H).
GN HNRPH1 OR HNRPH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 180-184; 193-197 AND 200-230.
RX MEDLINE=96081943; PubMed=7499401;
RA Honore B., Rasmussen H.H., Vorum H., Dejgaard K., Liu X.,
RA Gronov P., Madsen P., Gesser B., Tommerup N., Celis J.E.;
RT "Heterogeneous nuclear ribonucleoproteins H, H', and F are members of
RT a ubiquitously expressed subfamily of related but distinct proteins
RT encoded by genes mapping to different chromosomes.";
RL J. Biol. Chem. 270:28780-28789(1995).
RN [2]
RP SEQUENCE OF 127-135 AND 153-163.
RX MEDLINE=94203790; PubMed=7512260;
RA Matunis M.J., Xing J., Dreyfuss G.;
RT "The hnRNP F protein: unique primary structure, nucleic acid-binding
RT properties, and subcellular localization.";
RL Nucleic Acids Res. 22:1059-1067(1994).
RN [3]
RP SEQUENCE OF 200-230.
RX TISSUE-Keratinocytes;
RX MEDLINE=93162043; PubMed=1286667;
RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
RA Vandekerckhove J.;
RT "Microsequences of 145 proteins recorded in the two-dimensional gel
RT protein database of normal human epidermal keratinocytes.";
RL Electrophoresis 13:960-969(1992).
CC -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE HETEROGENEOUS NUCLEAR
CC RIBONUCLEOPROTEIN (HNRNP) COMPLEXES WHICH PROVIDE THE SUBSTRATE
CC FOR THE PROCESSING EVENTS THAT PRE-MRNAs UNDERGO BEFORE BECOMING
CC FUNCTIONAL. TRANSLATABLE MRNAs IN THE CYTOPLASM. BINDS POLY (RG).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC -1- TISSUE SPECIFICITY: EXPRESSED UBIQUITOUSLY.
CC -1- DOMAIN: EACH OVARY-RRM REPEAT BOUND POLY(RG), WHILE ONLY THE N-
CC TERMINAL ORRM BOUND POLY(RG) AND POLY (RU). NONE OF THE REPEATS
CC BOUND DETECTABLE AMOUNTS OF POLY(RA).

```

```

CC -1- SIMILARITY: CONTAINS 3 RNA RECOGNITION MOTIFS (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.lsb-sib.ch/announce/
CC or send an email to license@lsb.sib.ch).
CC -----
DR EMBL: L22009; AAA91346.1; -.
DR Aarhus/Chent-2DPAGE; 4410; IEF.
DR Aarhus/Chent-2DPAGE; 4429; IEF.
DR Aarhus/Chent-2DPAGE; 5416; IEF.
DR MIM: 601035; -.
DR InterPro: IPR000504; -.
DR Pfam: PF00076; rrm; 3.
DR PROSITE: PS50102; RRM; 3.
DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat.
FT DOMAIN 11 90 RNA-BINDING (RRM) 1.
FT DOMAIN 111 188 RNA-BINDING (RRM) 2.
FT DOMAIN 289 364 RNA-BINDING (RRM) 3.
FT DOMAIN 234 433 RNA-BINDING (RRM) 3.
FT REPEAT 234 249 1-1.
FT REPEAT 418 433 1-2.
FT DOMAIN 354 392 2 X 19 AA PERFECT REPEATS.
FT REPEAT 374 372 2-1.
FT REPEAT 374 392 2-2.
SQ SEQUENCE 449 AA; 49229 MW; 4ECF7A075C2526FF CRC64;

Query Match
Best Local Similarity 85.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIGHLYI 7
Db 175 RIGHRYI 181

RESULT 9
ID ROH2_HUMAN STANDARD; PRT; 449 AA.
AC P55795;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, last sequence update)
DT 01-OCT-2000 (Rel. 40, last annotation update)
DE HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN H' (HNRNP H') (FTP-3).
GN HNRPH2 OR FTP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95048329; PubMed=7959728;
RA Vorechovsky I., Vetric D., Holland J., Bentley D.R., Thomas K.,
RA Zhou J.N., Notarangelo L.D., Plebani A., Fontana G., Ochs H.D.;
RT "Isolation of cosmid and cDNA clones in the region surrounding the
RT BTK gene at Xq21.3-q22.";
RL Genomics 21:517-524(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Oeltjen J.C., Liu X., Lu J., Malley T.M., Allen R.C., Muzny D.M.,
RA Belmont J.W., Gibbs R.A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=96081943; PubMed=7499401;
RA Honore B., Rasmussen H.H., Vorum H., Dejgaard K., Liu X.,
RA Gronov P., Madsen P., Gesser B., Tommerup N., Celis J.E.;
RT "Heterogeneous nuclear ribonucleoproteins H, H', and F are members of

```

```

RT a ubiquitously expressed subfamily of related but distinct proteins
RL encoded by genes mapping to different chromosomes."
RU J. Biol. Chem. 270:28780-28789(1995).
CC -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE HETEROGENEOUS NUCLEAR
CC RIBONUCLEOPROTEIN (HNRP) COMPLEXES WHICH PROVIDE THE SUBSTRATE
CC FOR THE PROCESSING EVENTS THAT PRE-MRNAs UNDERGO BEFORE BECOMING
CC FUNCTIONAL, TRANSLATABLE MRNAs IN THE CYTOPLASM. BINDS POLY (RG).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOPLASM.
CC -1- TISSUE SPECIFICITY: EXPRESSED UNIPROTICALLY.
CC -1- SIMILARITY: CONTAINS 3 RNA RECOGNITION MOTIFS (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U01923; NOT_ANNOTATED_CDS.
DR EMBL: U78027; AAB64202.1; -.
DR Aarhus/Shent-2DPAGE: 4432; IEF.
DR MIM: 601036; -.
DR InterPro: IPR000504; -.
DR Pfam: PF00076; Rrm; 3.
DR PROSITE: PS50102; RRM; 3.
DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat.
FT DOMAIN 11 90 RNA-BINDING (RRM) 1.
FT DOMAIN 111 188 RNA-BINDING (RRM) 2.
FT DOMAIN 289 364 RNA-BINDING (RRM) 3.
FT DOMAIN 433 433 2 x 16 AA GLY-RICH APPROXIMATE REPEATS.
FT REPEAT 234 249 1-1.
FT REPEAT 418 433 1-2.
FT DOMAIN 354 392 2 x 19 AA PERFECT REPEATS.
FT REPEAT 354 372 2-1.
FT REPEAT 374 392 2-2.
SO SEQUENCE 449 AA; 49263 MW; C892523A638F07C7 CRC64;

Query Match 76.2%; Score 32; DB 1; Length 449;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RIGHTLY 7
DB 175 RIGHTLY 181

RESULT 10
C312_DROME STANDARD: PRT: 510 AA.
AC Q9YVN6;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROBABLE CYTOCHROME P450 312A1 (EC 1.14.-.-) (CYPOCCX11A1).
GN CYP312A1 OR CG5137.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

```

```

RA April J.F., Aghayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Bakendale J., Bayraktaroglu U., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Donnes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mlshina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Guan P., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Syrakas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: MAY BE INVOLVED IN THE METABOLISM OF INSECT HORMONES AND
CC -1- IN THE BREAKDOWN OF SYNTHETIC INSECTICIDES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
CC -1- OXIDIZED FLAVOPROTEIN + H(2)O.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM
CC (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE003521; AAF49275.1; -.
DR FlyBase: FBgn0036778; CYP312a1.
DR InterPro: IPR001128; -.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; FALSE_NEG.
KW Oxidoreductase; Monooxygenase; Membrane; Heme; Microsome;
FT Endoplasmic reticulum; Hypothetical protein.
FT BINDING 455 455 HEME (BY SIMILARITY).
SO SEQUENCE 510 AA; 58804 MW; 6973570BF91DE4E0 CRC64;

Query Match 76.2%; Score 32; DB 1; Length 510;
Best Local Similarity 85.7%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 IGHLYL 8
DB 43 IGHLYL 49

RESULT 11
MAG2_HUMAN STANDARD: PRT: 314 AA.
ID MAG2_HUMAN
AC P43356;

```

```

DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE MELANOMA-ASSOCIATED ANTIGEN 2 (MAGE-2 ANTIGEN).
GN MAGE2 OR MAGE2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=94102805; PubMed=8276455;
RA de Smet C., Lurquin C., van der Bruggen P., de Plaen E., Brasseur F.,
RA Boon F.;
RT "Sequence and expression pattern of the human MAGE2 gene.";
RL Immunogenetics 39:121-129(1994).
RN [2]
RN MUTAGENESIS.
RC TISSUE=Blood;
RX MEDLINE=94157413; PubMed=8113684;
RA Gaugler B., van den Eynde B., van der Bruggen P., Romero P.,
RA Caforio J.J., de Plaen E., Lethe B., Brasseur F., Boon T.;
RT "Human gene MAGE-3 codes for an antigen recognized on a melanoma by
RT autologous cytolytic T lymphocytes.";
RL J. Exp. Med. 179:921-930(1994).
CC -1- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC PROGRESSION. ANTIGEN RECOGNIZED ON A MELANOMA BY AUTOLOGOUS
CC CYTOLYTIC T LYMPHOCYTES.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES.
CC -1- SIMILARITY: BELONGS TO THE MAGE FAMILY. STRONG SIMILARITY (90%)
CC WITH MAGE12.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; L18920; AAA17729.1; -.
CC DR MIM; 300173; -.
CC DR InterPro; IPR002190; -.
CC DR Pfam; PF01454; MAGE; 1.
CC KW Antigen; Multigene family; Tumor antigen.
CC FT DOMAIN 40 43 POLY-SER.
CC FT MUTAGEN 170 170 V->D: IMPROVES ABILITY TO BIND TO HLA-A1.
CC SQ SEQUENCE 314 AA; 35055 MW; 844F16335A2BCE7 CRC64;

```

```

Query Match          73.8%; Score 31; DB 1; Length 314;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 2 IGHLYIL 8

DB 172 ISHLYIL 178

```

RESULT 12
MURC_BUCAL STANDARD; PRT; 484 AA.
AC P57310;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE UDP-N-ACETYLMURAMATE--ALANINE LIGASE (EC 6.3.2.8) (UDP-N-
DE ACETYLMURAMOYL-L-ALANINE SYNTHETASE).
GN MURC OR BU215.

```

```

OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RN SEQUENCE FROM N.A.
RX STRAIN=TOKYO 1998;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS.";
RL Nature 407:81-86(2000).
CC -1- FUNCTION: CELL WALL FORMATION (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + UDP-N-ACETYLMURAMOYL + L-ALANINE -
CC ADP + ORTHOPHOSPHATE + UDP-N-ACETYLMURAMOYL-L-ALANINE.
CC -1- PATHWAY: PEPTIDOGLYCAN BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE MURCDEF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AP001118; BAB2931.1; -.
CC KW Peptidoglycan synthetase; Cell wall; Cell division; Ligase;
CC KW ATP-binding.
CC FT NE_BIND 125 131 ATP (POTENTIAL).
CC SQ SEQUENCE 484 AA; 54311 MW; 06359DA35D5590A7 CRC64;

```

```

Query Match          73.8%; Score 31; DB 1; Length 484;
Best Local Similarity 71.4%; Pred. No. 57;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

OY 2 IGHLYIL 8

DB 268 IGHLYIL 274

```

RESULT 13
PARE_BORBU STANDARD; PRT; 599 AA.
AC 059189;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE TOPOISOMERASE IV SUBUNIT B (EC 5.99.1.1).
GN PARE OR BB0036.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RN SEQUENCE FROM N.A.
RX STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty J., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kierlavag A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gokeyne J.D., Weidman J.,
RA Uiterback T., Matthey L., McDonald L., Artach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).
RN [2]
RN SEQUENCE OF 1-83 FROM N.A.
RP STRAIN=212;

```

RX MEDLINE=95111614; PubMed=7812434;
 RA Ojalim C., Davidson B.E., Saint-Girons I., Old I.G.;
 RT Conservation of gene arrangement and an unusual organization of rRNA
 RT genes in the linear chromosomes of the Lyme disease spirochaetes
 RT *Borrelia burgdorferi*, B. garinii and B. afzelii.";
 RL Microbiology 140:2931-2940(1994).
 CC -1- FUNCTION: TOPOISOMERASE IV IS ESSENTIAL FOR CHROMOSOME
 CC SEGREGATION. IT HAS RELAXATION OF SUPERCOILED DNA ACTIVITY.
 CC PERFORMS THE DECATENATION EVENTS REQUIRED DURING THE REPLICATION
 CC OF A CIRCULAR DNA MOLECULE (BY SIMILARITY).
 CC -1- SUBUNIT: COMPOSED OF TWO SUBUNITS: PARC AND PARE.
 CC -1- SIMILARITY: BELONGS TO TYPE II TOPOISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE001117; AAC66418.1; -
 DR EMBL: L32861; AAC41408.1; -
 DR TIGR: BB0036; -
 DR InterPro: IPR001241; -
 DR Pfam: PF00204; DNA_topoisom; 1.
 DR PRINTS: PR00418; TP12FAMILY.
 DR PROSITE: PS00177; TOPOISOMERASE-IT; 1.
 KW Isomerase; Topoisomerase; ATP-binding;
 SQ SEQUENCE 599 AA; 68774 MW; B5901F7B1C7721 CRC64;

Query Match
 Best Local Similarity 73.8%; Score 31; DB 1; Length 599;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GHXYL 8
 Db 499 GHXYL 504

RESULT 14
 GCPE.CHLPN STANDARD; PRT; 613 AA.
 AC Q928H0: Q928H0 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DE GCPE PROTEIN HOMOLOG.
 GN GCPE OR CPN0373 OR CP0383.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiales; Chlamydiales; Chlamydiales.
 OX NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CW1029;
 RX MEDLINE=99206606; PubMed=10193388;
 RA Kalman S., Mitchell W., Marathe R., Lammell C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RL Nat. Genet. 21:385-389(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.R., Peterson J., Ullrich T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwyn M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 RT pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).

RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shiral M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CWL029 from USA.";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -1- SIMILARITY: BELONGS TO THE GCPE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE001621; AAD18517.1; -
 DR EMBL: AE002200; AAF38230.1; -
 DR EMBL: AP002546; BAA9581.1; -
 DR TIGR: CP0383; -
 SQ SEQUENCE 613 AA; 68308 MW; 40B1F831E02DC6D4 CRC64;

Query Match
 Best Local Similarity 73.8%; Score 31; DB 1; Length 613;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RIGHLY 7
 Db 19 RIGHLY 25

RESULT 15
 GYRB.BARBA STANDARD; PRT; 692 AA.
 AC P94281;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE DNA GYRASE SUBUNIT B (EC 5.99.1.3).
 GN GYRB.
 OS Bartonella bacilliformis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Bartonellaceae; Bartonella.
 OX NCBI_TaxID=774;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KC583 / ATCC 35685;
 RX MEDLINE=99013570; PubMed=9797224;
 RA Battisti J.M., Smithehman L.S., Samuels D.S., Minnick M.F.;
 RT "Mutations in Bartonella bacilliformis gyrb confer resistance to
 RT cotrimoxazole.";
 RL Antimicrob. Agents Chemother. 42:2906-2913(1998).
 CC -1- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-
 CC STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE
 CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
 CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
 CC -1- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING
 CC OF DOUBLE-STRANDED DNA.
 CC -1- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
 CC BREAKAGE AND REJOINING. THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
 CC ENZYME FORMS AN A2B2 TETRAMER.
 CC -1- SIMILARITY: BELONGS TO TYPE II TOPOISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@lsb-sib.ch).

DR EMBL; U82225; AAC71079.1; -.
DR InterPro; IPR001241; -.
DR InterPro; IPR002288; -.
DR Pfam; PF00986; DNA_gyraseB_C; 1.
DR Pfam; PF00204; DNA_topoisomII; 1.
DR PRINTS; PR00418; TP12FAMILY.
DR PROSITE; PS00177; TOPOISOMERASE_II; 1.
KW Topoisomerase; Isomerase; ATP-binding; Antibiotic resistance.
FT VARIANT 124 124 G -> S (IN COMMERMYCIN A1 MUTANT).
FT VARIANT 184 184 R -> Q (IN COMMERMYCIN A1 MUTANT).
FT VARIANT 214 214 T -> A (IN COMMERMYCIN A1 MUTANT).
FT VARIANT 214 214 T -> I (IN COMMERMYCIN A1 MUTANT).
SQ SEQUENCE 692 AA; 77445 MW; F54948EEF342281F CRC64;

Query Match 73.8%; Score 31; DB 1; Length 692;

Best Local Similarity 71.4%; Pred. No. 82;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIGHTLY 7
I I I I I
Db 577 RACHLYL 583

Search completed: June 20, 2001, 14:09:32
Job time: 329 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2001, 14:10:23 ; Search time 47.6 Seconds
(without alignments)
22.236 Million cell updates/sec

Title: US-09-692-401-6
Perfect score: 42
Sequence: 1 RIGHTLYL 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_PROTOZOA:*
12: SP_UNCLASSIFIED:*
13: SP_VERTEBRATE:*
14: SP_VIRUS:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	42	100.0	314 4 Q9NSD3	Q9nsd3 homo sapien
2	35	83.3	520 10 O65815	O65815 helianthus
3	35	83.3	1310 5 Q9VVF8	Q9vvf8 drosophila
4	34	81.0	234 4 Q9S1R5	Q9s1r5 streptomyce
5	34	81.0	1244 10 O82607	O82607 arabidopsis
6	34	81.0	1461 10 Q9S1M3	Q9s1m3 arabidopsis
7	33	78.6	305 2 O45454	O45454 bacillus su
8	33	78.6	324 2 Q924K0	Q924k0 streptomyce
9	33	78.6	444 2 O85748	O85748 streptomyce
10	33	78.6	723 4 Q9UDU6	Q9udus homo sapien
11	32	76.2	154 6 Q9WZ49	Q9wz49 ovys aries
12	32	76.2	297 4 Q9NP96	Q9np96 homo sapien
13	32	76.2	301 2 O52770	O52770 pseudomonas
14	32	76.2	331 4 Q9NPA7	Q9npa7 homo sapien
15	32	76.2	346 4 Q9Y4J5	Q9y4j5 homo sapien
16	32	76.2	350 2 Q9WZ92	Q9wz92 thermotoga
17	32	76.2	415 11 Q9Z2X1	Q9z2x1 rattus norv
18	32	76.2	449 11 Q35737	Q35737 mus musculu
19	32	76.2	449 11 P70333	P70333 mus musculu

20	32	76.2	657 4 Q9UJ28	Q9uj28 homo sapien
21	32	76.2	657 11 Q9ROC5	Q9roc5 rattus norv
22	32	76.2	666 5 Q9VZX5	Q9vzx5 drosophila
23	31	73.8	101 2 Q9PH51	Q9ph51 xylenella fas
24	31	73.8	139 2 O49122	O49122 methylobact
25	31	73.8	146 14 Q9J511	Q9j511 fowlpox vir
26	31	73.8	170 4 Q9NPM4	Q9npm4 homo sapien
27	31	73.8	176 2 O87064	O87064 vibrio chol
28	31	73.8	178 11 O35486	O35486 mus musculu
29	31	73.8	181 4 Q9NPP3	Q9npp3 homo sapien
30	31	73.8	204 13 Q9PRU7	Q9pru7 oreochromis
31	31	73.8	232 14 Q9Q22	Q9q22 gallid hefp
32	31	73.8	233 10 Q9S181	Q9s181 arabidopsis
33	31	73.8	269 8 Q9LW9	Q9lwn9 kluyveromyc
34	31	73.8	276 4 Q9Y4M4	Q9y4m4 homo sapien
35	31	73.8	290 4 Q9NPS3	Q9nps3 homo sapien
36	31	73.8	336 5 Q9VAT5	Q9vat5 drosophila
37	31	73.8	339 4 Q9H8E1	Q9h8e1 homo sapien
38	31	73.8	378 5 Q9VRC3	Q9vrc3 drosophila
39	31	73.8	423 10 O49639	O49639 arabidopsis
40	31	73.8	465 2 O86374	O86374 mycobacteri
41	31	73.8	479 11 Q9EP14	Q9ep14 rattus norv
42	31	73.8	497 11 O35173	O35173 mus musculu
43	31	73.8	497 11 O88758	O88758 rattus norv
44	31	73.8	509 11 Q9ER21	Q9er21 mus musculu
45	31	73.8	509 11 Q9EP15	Q9ep15 rattus norv

ALIGNMENTS

```

RESULT 1
Q9NSD3 PRELIMINARY; PRT; 314 AA.
AC Q9NSD3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE MELANOMA ANTIGEN FAMILY A12.
GN MAGP12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Mallon A.M., Platzer M., Bates R., Gloeckner G., Botcherby M.,
RA Nordstiek G., Strivens M.A., Kioschis P., Dangel A., Cunningham D.,
RA Straw R., Weston P., Hunter C., Gilbert M., Fernando S., Goodall K.,
RA Kerry G., Greystrong J.S., Clark D., Goerdes M., Blechschmidt K.,
RA Rump A., Hinzmann B., Mundy C.R., Miller W., Foustka A., Herman G.E.,
RA Rhodes M., Denny P., Rosenthal A., Brown S.D.M.;
RT "Comparative genome sequence analysis of the Bpa/Str region in mouse
RT and man.";
RL Genome Res. 0:0-0(2000).
DR EMBL; U82671; AAF44789.1; -.
DR InterPro: IPR002190; -.
DR Pfam; PF01454; MAGP. 1.
SO SEQUENCE 314 AA; 34836 MW; 7E0DF7CECD8F6568 CRC64;

Query Match 100.0%; Score 42; DB 4; Length 314;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RIGHTLYL 8
DB 171 RIGHTLYL 178

RESULT 2
Q65815 PRELIMINARY; PRT; 520 AA.
ID 065815

```

AC 065815: 065814:
 DT 01-AUG-1998 (Tremblrel. 07, last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, last annotation update)
 DE CYTOCHROME P450 8JBI (EC 1.14.13.53) (ISOFLAVONE 2'-HYDROXYLASE).
 GN CYP81B1 OR CYP81B1 OR CYP81B1C OR CYP81B1S.
 OS Helianthus tuberosus (Jerusalem artichoke).
 OC Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:
 OC Magnoliophyta: eudicotyledons, core eudicots: Asteridae:
 OC Eriasterids II; Asterales; Asteraceae; Asteroideae; Heliantheae;
 OC Helianthus.
 OX NCBI_TaxID=4233;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. BLANC COMMUN; TISSUE=TUBER;
 RX MEDLINE=98184826; PubMed=9516419;
 RA Cabello-Hurtado F., Batard Y., Saleun J.-P., Durst F., Pinot F.,
 RA Werck-Reichardt D.;
 RT "Cloning, expression in yeast, and functional characterization of
 RT CYP81B1, a plant cytochrome P450 that catalyzes in-chain hydroxylation
 RT of fatty acids.";
 RL J. Biol. Chem. 273:7260-7267(1998).
 CC
 CC -1- FUNCTION: CATALYZES THE HYDROXYLATION OF ISOFLAVONES, DAIDZEIN AND
 CC FORMONONETIN, TO YIELD 2'-HYDROXYISOFLAVONES, 2'-HYDROXYDAIDZEIN,
 CC AND 2'-HYDROXYFORMONONETIN, RESPECTIVELY (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: FORMONONETIN + NADPH + O(2) =
 CC 2-HYDROXYFORMONONETIN + NADP(+) + H(2)O.
 CC -1- COFACTOR: THE ENZYME IS NADPH-DEPENDENT.
 CC -1- ENZYME REGULATION: ENZYME ACTIVITY IS INDUCED BY MNCL2 OR
 CC AMINOPURINE.
 CC -1- PATHWAY: INVOLVED IN THE BIOSYNTHESIS OF ISOFLAVONOID-DERIVED
 CC ANTIMICROBIAL COMPOUNDS OF LEGUMES.
 CC -1- ALTERNATIVE PRODUCTS: TWO FORMS: CYP81B1 (SHOWN HERE) AND
 CC CYP81B1S; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL: AJ000478; CA04117.1; -;
 DR EMBL: AJ000477; CA04116.1; -;
 DR InterPro: IPR001128; -;
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
 DR Oxidoreductase; Monooxygenase; Membrane; Heme; Alternative splicing;
 KW NADP.
 FT BINDING 456 456 HEME (BY SIMILARITY).
 FT VAAPSLIC 9 22 MISSING (IN CYP81B1S).
 FT CONFLICT 40 40 MISSING (IN CA04116).
 FT CONFLICT 51 51 Y -> S (IN CA04116).
 FT CONFLICT 78 78 P -> Q (IN CA04116).
 SQ SEQUENCE 520 AA; 58914 MW; 38108E379BDB35C8 CRC64;

Query Match 83.3%; Score 35; DB 10; Length 520;
 Best Local Similarity 85.7%; Pred. No. 57;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 IGHLYIL 8
 |||||:|
 Db 55 IGHLYIL 61

RESULT 3
 Q9VVF8 PRELIMINARY; PRT; 1310 AA.
 AC Q9VVF8;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
 DE 01-MAY-2000 (Tremblrel. 13, last annotation update)
 DE CG7692. PROTEIN.
 GN CG7692.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta;
 OC Pterygota: Neoptera: Endopterygota: Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amarantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pianko C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Boltskov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Peltan G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reiner K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003524; AAF49353.1; -;
 DR FlyBase: FBgn0036714; CG7692.
 SQ SEQUENCE 1310 AA; 150503 MW; 2AF0754E8452430 CRC64;

Query Match 83.3%; Score 35; DB 5; Length 1310;
 Best Local Similarity 85.7%; Pred. No. 1,5e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 IGHLYIL 8
 |||||:|
 Db 219 IGHLYIL 225

RESULT 4
 Q9S1R5 PRELIMINARY; PRT; 234 AA.
 AC Q9S1R5;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
 DE 01-MAY-2000 (Tremblrel. 13, last annotation update)
 DE PUTATIVE MEMBRANE PROTEIN.
 GN SCJ9A.05C.
 OS Streptomyces coelicolor
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);

RA Harris D.;
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Ralmandream M.A.;
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RX MEDLINE-97000351; PubMed-8843436;
RA Kadenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,
Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RT Mol. Microbiol. 21:77-96(1996).
DR EMBL: AL109972; CAB5266.1; -;
SQ SEQUENCE 234 AA; 25619 MW; AAFEB1CD014319 CRC64;

Query Match 81.0%; Score 34; DB 2; Length 234;
Best Local Similarity 71.4%; Pred. No. 39;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 RIGHLYTL 8
DB 150 RIGHLYTL 156

RESULT 5
ID 082607 PRELIMINARY; PRT; 1244 AA.
AC 082607;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE T2L5.9 PROTEIN.
GN T2L5.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA WASHU;
RT "The A. thaliana Genome Sequencing Project.";
RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Geisel C., Smith A., Le T.;
RT "The sequence of A. thaliana T2L5.";
RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Waterston R.;
RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: BELONGS TO ZN-FINGER CCHC TYPE FAMILY.
DR EMBL: AF096371; AAC62795.1; -;
DR InterPro: IPR001584; -;
DR InterPro: IPR001878; -;
DR Pfam: PF00098; zf-CCHC; 1.
DR Pfam: PF00665; tve; 1.
DR PRINTS: PR00939; C2HCZNFINGER.
DR SMART: SM00343; znf_C2HC; 1.
KW zinc-finger.
SQ SEQUENCE 1244 AA; 139337 MW; C2B34AA509B5C80 CRC64;

Query Match 81.0%; Score 34; DB 10; Length 1244;

Best Local Similarity 75.0%; Pred. No. 2,3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RIGHLYTL 8
DB 520 RIGHLYTL 527

RESULT 6
ID 09S1M3 PRELIMINARY; PRT; 1461 AA.
AC 09S1M3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PUTATIVE RETROELEMENT POL POLIPROTEIN.
GN ATG20460.
GN Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RX MEDLINE-20083487; PubMed-10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
Fuji C.T., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome II of Arabidopsis thaliana.";
RL Nature 402:761-768(1999).
CC -I- SIMILARITY: BELONGS TO ZN-FINGER CCHC TYPE FAMILY.
DR EMBL: AC007109; AAD25646.1; -;
DR InterPro: IPR001584; -;
DR InterPro: IPR001878; -;
DR Pfam: PF00098; zf-CCHC; 1.
DR Pfam: PF00665; tve; 1.
DR SMART: SM00343; znf_C2HC; 1.
KW Polyprotein; zinc-finger.
SQ SEQUENCE 1461 AA; 162670 MW; E79E2C46BBEE3D16 CRC64;

Query Match 81.0%; Score 34; DB 10; Length 1461;
Best Local Similarity 75.0%; Pred. No. 2,7e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RIGHLYTL 8
DB 544 RIGHLYTL 551

RESULT 7
ID 045454 PRELIMINARY; PRT; 305 AA.
AC 045454;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE PROBABLY POSITIONED IN OPERON WITH DOWNSTREAM ORFs.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-IF03022;
RX MEDLINE-95206941; PubMed-7899081;
RA Meijer W.J., Venema G., Bron S.;

RT "Characterization of single strand origins of cryptic rolling-circle
RT plasmids from *Bacillus subtilis*.";
RL Nucleic Acids Res. 23:612-619(1995).
DR EMBL: U32380; AAC44422.1; -.
KW Plasmid.
SQ SEQUENCE 305 AA; 35592 MW; DB410B2A7BC51B22 CRC64;

Query Match 78.6%; Score 33; DB 2; Length 305;
Best Local Similarity 71.4%; Pred. No. 82;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 IGHLYIL 8
|||:|:|
DB 199 IGHYYL 205

RESULT 8
O924K0 PRELIMINARY; PRT; 324 AA.
AC O924K0:
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE FACTOR C PROTEIN PRECURSOR.
OS Streptomyces griseus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=45H;
RX MEDLINE=80156796; PubMed=6767606;
RA Biro S., Bekesi I., Vitalis S., Szabo G.;
RT "A substance effecting differentiation in *Streptomyces griseus*.
RT Purification and properties.";
RL Eur. J. Biochem. 103:359-363(1980).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=45H;
RA Birko Z., Simegl A., Szeszak F., Vitalis S., Szabo G., Biro S.;
RT "Cloning and sequencing of the factor C gene from *Streptomyces griseus*
RT 45H.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF103943; AAC97368.1; -.
KW Signal.
FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 324 FACTOR C PROTEIN.
SQ SEQUENCE 324 AA; 34555 MW; E5F57BA09D2B4EDB CRC64;

Query Match 78.6%; Score 33; DB 2; Length 324;
Best Local Similarity 50.0%; Pred. No. 88;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 RIGHLYIL 8
::|:|:|
DB 106 RIGHMYLL 113

RESULT 9
O85748 PRELIMINARY; PRT; 444 AA.
AC O85748:
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE CHEW.
GN CHEW.
OS *Treponema denticola*.
OC Bacteria; Spirochaetales; Spirochaetaceae; *Treponema*.
OX NCBI_TaxID=158;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=ATCC35405;
RA Greene S.R., Stamm L.V.;
RT "Treponema denticola chemotaxis genes.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF074950; AAC33469.1; -.
DR InterPro: IPR000780; -.
DR InterPro: IPR001601; -.
DR InterPro: IPR002545; -.
DR Pfam: PF01584; Chew; 1.
DR Pfam: PF01739; Cher; 1.
DR SMART: SM00138; Metrc; 1.
SQ SEQUENCE 444 AA; 49706 MW; 2C1CE0F6A1D6108E CRC64;

Query Match 78.6%; Score 33; DB 2; Length 444;
Best Local Similarity 100.0%; Pred. No. 1,2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 GHLVIL 8
|||||
DB 154 GHLVIL 159

RESULT 10
O9UD06 PRELIMINARY; PRT; 723 AA.
AC O9UD06:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE HEPATOCYTE GROWTH FACTOR.
GN HGF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Courtney L., Elliot G., Angel S.;
RT "The sequence of Homo sapiens PAC clone RP5-1098B1.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Waterston R.H.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC -!- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR EMBL: AC004960; AAC71655.1; -.
DR HSSP: P14210; IBHT.
DR InterPro: IPR000001; -.
DR InterPro: IPR000327; -.
DR InterPro: IPR001254; -.
DR InterPro: IPR001314; -.
DR InterPro: IPR003014; -.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00051; Kingle; 4.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000583; -. 1.
DR PROSITE: PS00021; KRINGLE_1; UNKNOWN_4.
DR PROSITE: PS0070; KRINGLE_2; 4.
DR SMART: SM00020; Tryp_SPC; 1.
KW Hydrolyase; Serine protease.
SQ SEQUENCE 723 AA; 82602 MW; 627B1EF99FAD931B CRC64;

Query Match 78.6%; Score 33; DB 4; Length 723;

Best Local Similarity 62.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIGHTLY 8
DB 625 RVHLYTM 632

RESULT 11

Q9MZA9 PRELIMINARY; PRT; 154 AA.
AC Q9MZA9; 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE VIMENTIN (FRAGMENT).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTAL ARTERY;
RA Zheng J., Tsol S.C., Magness R.R.;
RT "Growth factor expression in ovine fetal placental artery endothelial cells."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF251147; AAF87226.1; -
DR InterPro: IPR001664; -
DR Pfam: PF00038; Filament; 1.
FT NON_TER 154
SQ SEQUENCE 154 AA; 17171 MW; 3C1632E79DE31A49 CRC64;

Query Match 76.2%; Score 32; DB 6; Length 154;
Best Local Similarity 83.3%; Pred. No. 63;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIGHTLY 6
DB 145 RLGHLY 150

RESULT 12
Q9NP96 PRELIMINARY; PRT; 297 AA.
AC Q9NP96; 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE HNRNP 2H9B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20461765; PubMed=10858537;
RA Honore B.;
RT "The hnrnp 2H9 gene, which is involved in the splicing reaction, is a multiply spliced gene."
RL Blochim. Biophys. Acta 1492:108-119(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Honore B.;
RT "The hnrnp 2H9 gene, which is involved in the splicing reaction, is a multiply spliced gene."
RL Blochim. Biophys. Acta 0:0-0(2000).
DR EMBL: AF132362; AAF68850.1; -
DR EMBL: AF132360; AAF68844.1; -
DR InterPro: IPR000504; -
DR Pfam: PF00076; rrm; 1.

SQ SEQUENCE 297 AA; 31525 MW; AAE63A29063E5153 CRC64;

Query Match 76.2%; Score 32; DB 4; Length 297;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIGHTLY 7
DB 31 RIGHRYI 37

RESULT 13

Q52770 PRELIMINARY; PRT; 301 AA.
AC Q52770; 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE ACTIVATOR PROTEIN.
GN ACTR.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 50106;
RA Bruenker P., Altenbuchner J.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: BELONGS TO THE ARAC/XYS FAMILY OF TRANSCRIPTIONAL REGULATORS.
DR EMBL: AF047527; AAC34292.1; -
DR InterPro: IPR000005; -
DR Pfam: PF00165; HTH_ARAC; 1.
DR PRINTS: PR00032; HTHARAC.
DR PROSITE: PS00041; HTH_ARAC_FAMILY_1; 1.
DR PROSITE: PS01124; HTH_ARAC_FAMILY_2; 1.
DR SMART: SM00342; HTH_ARAC; 1.
KW DNA-binding; Transcription regulation.
SQ SEQUENCE 301 AA; 34776 MW; 08D13FAF5524DD06 CRC64;

Query Match 76.2%; Score 32; DB 2; Length 301;
Best Local Similarity 62.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RIGHTLY 8
DB 154 RLGHFFIL 161

RESULT 14
Q9NP97 PRELIMINARY; PRT; 331 AA.
AC Q9NP97; 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE HNRNP 2H9A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20461765; PubMed=10858537;
RA Honore B.;
RT "The hnrnp 2H9 gene, which is involved in the splicing reaction, is a multiply spliced gene."
RL Blochim. Biophys. Acta 1492:108-119(2000).
RN [2]
RP SEQUENCE FROM N.A.

RA Honore B.;
 RT "The hnRNP 2H9 gene, which is involved in the splicing reaction, is a
 RT multiply spliced gene.";
 RL Biochim. Biophys. Acta 0:0-0(2000).
 DR EMBL: AF132361; AAF68849.1; -.
 DR EMBL: AF132360; AAF68847.1; -.
 DR InterPro: IPR000504; -.
 DR Pfam: PF00076; ttm: 2.
 SO SEQUENCE 331 AA; 35239 MW; A3BA8D05FEC07369 CRC64;

Query Match 76.2%; Score 32; DB 4; Length 331;
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RIGHLYI 7
 |||||
 Db 80 RIGHRYI 86

RESULT 15
 O9Y4J5 PRELIMINARY; PRT; 346 AA.
 ID O9Y4J5
 AC O9Y4J5;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE RIBONUCLEOPROTEIN (HNRNP 2H9).
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97153147; PubMed=8999868;
 RA Mahe D., Mahl P., Gattoni R., Fischer N., Mattei M.G., Stevenin J.,
 RA Fuchs J.P.;
 RT "Cloning of human 2H9 heterogeneous nuclear ribonucleoproteins.
 RT Relation with splicing and early heat shock-induced splicing arrest.";
 RL J. Biol. Chem. 272:1827-1836(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Honore B.;
 RT "The hnRNP 2H9 gene, which is involved in the splicing reaction, is a
 RT multiply spliced gene.";
 RL Biochim. Biophys. Acta 0:0-0(2000).
 DR EMBL: L32610; AADA5179.1; -.
 DR EMBL: AF132360; AAF68843.1; -.
 DR InterPro: IPR000504; -.
 DR Pfam: PF00076; ttm: 2.
 KW Nucleocapsid; Ribonucleoprotein.
 SO SEQUENCE 346 AA; 36926 MW; F7D14C2947930E9E CRC64;

Query Match 76.2%; Score 32; DB 4; Length 346;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RIGHLYI 7
 |||||
 Db 80 RIGHRYI 86

Search completed: June 20, 2001, 14:10:24
 Job time: 346 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2001, 14:04:02 ; Search time 41.58 Seconds

(Without alignments)
10.998 Million cell updates/sec

Title: US-09-692-401-6

Perfect score: 42

Sequence: 1 RIGHLYL 8

Scoring table: BLOSUM62

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

1: /SID56/gcgdata/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	83.3	505	20	Y05902
2	34	81.0	9	15	R50283
3	34	81.0	9	17	R93435
4	34	81.0	9	20	Y46062
5	34	81.0	10	18	W23043
6	33	78.6	9	20	Y46263
7	33	78.6	9	20	Y47329
8	33	78.6	9	20	Y47950
9	33	78.6	10	20	Y46214
10	33	78.6	10	20	Y46333
11	33	78.6	10	20	Y47356

12.	33	78.6	10	20	Y47357	Immunogenic peptid
13	33	78.6	10	20	Y48011	Immunogenic peptid
14	33	78.6	15	21	B08833	Amino acid sequenc
15	33	78.6	15	21	Y98994	H1A class II bindl
16	33	78.6	125	18	W16331	Baboon MAG-3 homo
17	33	78.6	126	18	W16330	Human MAG-3 tumou
18	33	78.6	189	13	R25592	Human MAG-3-His f
19	33	78.6	314	20	Y01720	Exons XV to XVIIII
20	33	78.6	314	20	B02565	The MAG-3 polypep
21	33	78.6	403	20	Y06591	Human MAG-A3 prot
22	33	78.6	451	20	Y06589	Haemagglutinin-MAG
23	33	78.6	453	20	Y06593	Lipoprotein D-MAGE
24	33	78.6	697	21	Y98485	CLYTR-MAG-3-His f
25	33	78.6	697	21	Y59030	Hepatocyte growth
26	33	78.6	697	22	B45838	Sequence of a pept
27	33	78.6	723	11	R07144	Nucleic acid trans
28	33	78.6	723	12	R14307	Tumour cytotoxic f
29	33	78.6	723	12	R15624	Plasminogen-like g
30	33	78.6	723	13	R21142	Human leukocyte-de
31	33	78.6	723	13	R25677	Human TCF-II. Hom
32	33	78.6	723	13	R29819	Recombinant human
33	33	78.6	723	15	R57026	TCF-II. R29819;
34	33	78.6	723	15	R57027	Human wild-type tu
35	33	78.6	723	15	R57028	Human modified tum
36	33	78.6	723	15	R82685	Human modified tum
37	33	78.6	723	17	R96688	Tumour cytotoxic f
38	33	78.6	723	17	R96689	TCF mutant having
39	33	78.6	723	19	W76690	TCF mutant having
40	33	78.6	723	19	W59923	Human plasmidogen-
41	33	78.6	723	12	R10656	Human leukocyte-de
42	33	78.6	728	12	R12792	Hepatic parenchyma
43	33	78.6	728	12	R14243	Human hepatocyte g
44	33	78.6	728	12	R15623	Human hepatocyte g
45	33	78.6	728	13	R20005	Human leukocyte-de
						Human hepatocyte g

ALIGNMENTS

RESULT 1	Y05902	standard; Protein, 505 AA.
XX	Y05902;	
AC	XX	
DT	XX	
DE	XX	02-AUG-1999 (first entry)
XX	XX	
XX	XX	Jerusalem artichoke in-chain hydroxylase CYP81B1.
XX	XX	In-chain hydroxylase; transgenic plant; lipid; hydroxylation;
KW	KW	oilseed; vegetable oil; crop protection; Jerusalem artichoke;
KW	KW	CYP81B1; cytochrome P450.
XX	XX	
OS	XX	Helianthus tuberosus.
PN	XX	WO9918224-A1.
XX	XX	
PD	XX	15-APR-1999.
PE	XX	06-OCT-1998; 98WO-IB01716.
PR	XX	06-OCT-1997; 97US-0060960.
XX	XX	
XX	XX	(CNRS) CENT NAT RECH SCI.
PI	PI	Batard Y, Benveniste I, Cabello-Huartado F, Durst F;
PI	PI	Heivig C, Le Bouquin R, Pinot F, Salaun J, Tijet N;
XX	XX	Merck-Reichart D;
DR	DR	WPI: 1999-264030/22.
XX	XX	N-PSDB; X58406.
PT	PT	Nucleic acid encoding plant fatty acid hydroxylases

XX PS Example 4; Fig 20A-B; 157pp; English.

CC CC The present sequence represents in-chain hydroxylase CYP81B1 of

CC CC Jerusalem artichoke. CYP81B1 is a microsomal cytochrome P450 that

CC CC catalyses the omega-2, omega-3 and omega-4 hydroxylation of capric,

CC CC lauric and myristic acids. The major metabolite is the

CC CC omega-3-hydroxylated compound. The invention provides isolated

CC CC nucleic acids (see X58400-06) encoding plant fatty acid

CC CC hydroxylases (see Y05896-902). Also claimed are host cells,

CC CC transgenic plants and compositions consisting of the plant fatty

CC CC acid hydroxylase, a process for isolating additional fatty acid

CC CC hydroxylase genes from a plant, and a process of altering the fatty

CC CC acid composition in a plant by expressing the plant fatty acid

CC CC hydroxylase in a transgenic plant, and hydroxylating or epoxidating

CC CC a fatty acid substrate in the plant. Manipulating the hydroxylated

CC CC fatty acid content of plants will modify resistance to drought and

CC CC attack by insects and other pests. The transgenic plants may also

CC CC be used as sources of hydroxylated and epoxidized fatty acids

CC CC useful in the manufacture of e.g. lubricants, anti-slip agents,

CC CC plasticisers, coating agents, detergents and surfactants.

CC XX

SO Sequence 505 AA;

Query Match 83.3%; Score 35; DB 20; Length 505;

Best Local Similarity 85.7%; Pred. No. 24;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 IGHLYIL 8

|||||

Db 40 Ighlyl 46

RESULT 2

ID R50283 standard; Protein; 9 AA.

XX AC R50283;

XX DT 26-SEP-1994 (first entry)

XX DE MAGE-21 nonapeptide.

XX KM MAGE: nonapeptide; cancer; melanoma; breast cancer; HLA;

XX KM histocompatibility; human leucocyte antigen; probe; treatment;

XX KM therapy; vaccine.

XX OS Synthetic.

XX PN WO9405304-A.

XX PD 17-MAR-1994.

XX PF 30-AUG-1993; 93WO-US08157.

XX PR 31-AUG-1992; 92US-0938334.

XX PR 26-MAR-1993; 93US-0037230.

XX PR 07-JUN-1993; 93US-0073103.

XX PA (LUDW-) LUDWIG INST CANCER RES.

XX PI Boon-fallieur T, De Plaen E, Lurquin C, Traversari C;

XX PI Van Derbruggen P;

XX DR WPI: 1994-100844/12.

XX DR N-PSDB: Q44753.

XX PT New nona-peptide derived from tumour rejection antigen precursor

XX PT - presented by HLA-A1 cancer cells, for use in diagnosis or

XX PT therapy of esp. melanoma and breast cancer.

XX PS Disclosure; Page 20; 33pp; English.

XX CC An isolated nonapeptide having the amino acid sequence Glu-Val-Asp-

CC CC pro-tle-gly-His-Ileu-Tyr is derived from the tumour rejection antigen

CC CC precursor encoded by the MAGE-3 gene and presented by HLA-A1. The

CC CC nonapeptide can be used in a vaccine to treat a cancerous condition

CC CC involving HLA-A1 subtype cancerous cells. The nucleic acid encoding

CC CC the nonapeptide can be used as a probe to identify tumour cells.

CC CC This sequence is homologous to the peptide described and is encoded

CC CC by the MAGE-21 gene.

CC XX

SO Sequence 9 AA;

Query Match 81.0%; Score 34; DB 15; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.2e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RIGHLY 6

|||||

Db 4 righly 9

RESULT 3

ID R99345 standard; Protein; 9 AA.

XX AC R99345;

XX DT 22-APR-1997 (first entry)

XX DE MAGE-21 nonapeptide.

XX KM HLA binding peptide; cell lysis; cytolytic T cell; MAGE family; human;

XX KM tumour rejection antigen precursor; TRA; MAGE-1; tumour; cancer cell;

XX KM antibody; melanoma; universal effector cell; vaccine; breast cancer; CTL;

XX KM therapy.

XX OS Homo sapiens.

XX PN WO9626214-A1.

XX PD 29-AUG-1996.

XX PF 01-FEB-1996; 96WO-US01489.

XX PR 23-FEB-1995; 95US-0393273.

XX PA (LUDW-) LUDWIG INST CANCER RES.

XX PI Boon-Fallieur T, De Plaen E, Gaugler B, Lurquin C;

XX PI Romero P, Traversari C, Van Den Eynde B, Van Der Bruggen P;

XX DR WPI: 1996-402317/40.

XX DR N-PSDB: T35410.

XX PT New nona-peptide(s) that bind to HLA molecule(s) and induce lysis

XX PT by specific cytolytic T cells, for diagnosis and treatment of

XX PT tumours and to expand T cells in vitro.

XX PS Example 4; Fig 4; 41pp; English.

CC CC R99343-R99350 represent MAGE nonapeptides, based on the tumour rejection

CC CC antigen region of the full length MAGE sequences. These peptides were

CC CC used to design the nonapeptides of the invention (see R99337-R99342),

CC CC which bind to a HLA molecule on a cell, and provoke lysis by cytolytic T

CC CC cells (CTLs) specific for a complex of the HLA molecule and nonapeptide.

CC CC The nonapeptides can be used diagnostically to identify tumours

CC CC expressing a particular HLA molecule, or to identify cancer cells. The

CC CC peptides can also be used therapeutically, to induce a CTL response to

CC CC tumours (where the peptides are optionally coupled to tumour-specific

CC CC antibodies), or to induce a response by CTLs that are otherwise inactive.

CC CC The peptide sequences may also be used to expand specific CTLs in vitro

CC CC for later return to the patient, such as for treating melanoma. Tumour

CC cells can be identified by using DNA encoding the nonapeptides as probes.
 CC Non-human cells transformed with the HLA-A1 gene and a DNA sequence
 CC encoding one of the peptides, can be used to generate CTLs or to detect
 CC the presence of CTLs in human samples. The non-human transformed cells,
 CC when polyclonally transformed, are universal effector cells, and can be used in
 CC vaccines, or for treating melanoma or breast cancer.

CC Sequence 9 AA;

Query Match 81.0%; Score 34; DB 17; Length 9;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIGHTLY 6
 |||||
 Db 4 rightly 9

RESULT 4

Y46062 Y46062 standard; Peptide: 9 AA.

AC Y46062;

DT 01-DEC-1999 (first entry)

DE Immunogenic peptide having a human leukocyte antigen binding motif #673.

XX Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
 KW Immune response; T cell activation; major histocompatibility complex;
 KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
 KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
 KW vaccine; immunisation.

XX Synthetic.

OS Homo sapiens.

PN W09945954-A1.

PD 16-SEP-1999.

PF 13-MAR-1998; 98WO-US05039.

PR 13-MAR-1998; 98WO-US05039.

PA (EPIM-) EPIMUNE INC.

PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;

DR WPI; 1999-551214/46.

PT New immunogenic peptides with HLA binding motif, useful in treatment
 and diagnosis of cancers and viral diseases -

PS Claim 1; Page 54; 150pp; English.

XX Y45390 to Y48214 represent specifically claimed immunogenic peptides
 CC having a human major histocompatibility complex (MHC) Class I (also
 CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
 CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
 CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
 CC response against the antigen from which the peptide is derived.
 CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
 CC normally induced by an antigen in the form of a peptide fragment bound
 CC to a HLA molecule, rather than the intact foreign antigen itself, and
 CC are particularly important in tumour rejection and in fighting viral
 CC infections. The peptides are therefore useful therapeutically to treat
 CC or prevent viral infections and cancers in mammals (especially humans)
 CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
 CC They can be administered as vaccines to elicit an immune response in
 CC individuals susceptible or otherwise at risk of viral infection or
 CC cancer, or used to treat chronic or acute conditions. They are also

CC useful diagnostically, and can be used to induce a cytotoxic T cell
 CC response, by contacting a cytotoxic T cell with the peptide e.g. to
 CC produce CTLs ex vivo for infusion back into a patient. The
 CC polynucleotides encoding the immunogenic peptides are also useful
 CC therapeutically and for immunisation as above.

CC Sequence 9 AA;

Query Match 81.0%; Score 34; DB 20; Length 9;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RIGHTLY 6
 |||||
 Db 4 rightly 9

RESULT 5

W23043 W23043 standard; Peptide: 10 AA.

AC W23043;

DT 25-FEB-1998 (first entry)

DE MAGE-12/HLA-B44 tumour rejection antigen.

XX MAGE-12; tumour rejection antigen precursor; TRAP; HLA-B44;
 KW human leukocyte antigen B44; cytotoxic T lymphocyte; cancer;
 KW melanoma; therapy; diagnosis; vaccine.
 XX Homo sapiens.

PN W09731017-A1.

PD 28-AUG-1997.

PF 05-FEB-1997; 97WO-US01915.

PR 20-FEB-1996; 96US-0602506.

PA (LUDW-) LUDWIG INST CANCER RES.

PI Boon-Falleur T, Coulie P, Herman J, Luescher I;

DR WPI; 1997-435086/40.

PT Tumour rejection antigens presented by human leukocyte antigen B44
 molecules - useful to identify HLA-B44 positive cells for diagnosis
 and therapy of cellular abnormalities

PS Claim 2; Page 51; 74pp; English.

XX This peptide is a tumour rejection antigen presented by a HLA-B44
 CC molecule and derived from a MAGE-12 tumour rejection antigen
 CC precursor (TRAP). Claimed tumour rejection antigens (W23038-43)
 CC are able to bind to HLA-B44 positive cells, making them useful in
 CC identifying cells which present HLA-B44 molecules on their
 CC surfaces for use in the diagnosis and therapy of cellular
 CC abnormalities. The complex of the tumour rejection antigen and HLA
 CC molecule provokes a cytolytic T cell response. The tumour
 CC rejection antigens, or complexes of tumour rejection antigens and
 CC HLA-B44, can be used as vaccines to treat disorders characterised
 CC by expression of the TRAP molecule such as cancer, especially
 CC melanoma. Vaccines can also be prepared from cells which present
 CC the tumour rejection antigen/HLA complexes on their surface, such
 CC as non-proliferative cancer cells and non-proliferative
 CC transfectants.

XX Sequence 10 AA;

Query Match 81.0%; Score 34; DB 18; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.56;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGHLY 6
 |||||
 Db 5 ighly 10

RESULT 6
 Y46263
 ID Y46263 standard; Peptide; 9 AA.
 XX Y46263;
 AC
 XX
 DE 01-DEC-1999 (first entry)
 XX
 DE Immunogenic peptide having a human leukocyte antigen binding motif #874.
 XX
 KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
 KW immune response; T cell activation; major histocompatibility complex;
 KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
 KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
 KW vaccine; immunisation.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9945954-A1.
 XX
 PD 16-SEP-1999.
 XX
 PF 13-MAR-1998; 98WO-US05039.
 XX
 PR 13-MAR-1998; 98WO-US05039.
 XX
 PA (EPIM-) EPIMUNE INC.
 XX
 PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
 XX
 DR WPI; 1999-551214/46.
 XX
 PT New immunogenic peptides with HLA binding motif, useful in treatment
 PT and diagnosis of cancers and viral diseases -
 XX
 PS Claim 1; Page 64; 150pp; English.
 XX
 CC Y45390 to Y48214 represent specifically claimed immunogenic peptides
 CC having a human major histocompatibility complex (MHC) Class I (also
 CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
 CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
 CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
 CC response against the antigen from which the peptide is derived.
 CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
 CC normally induced by an antigen in the form of a peptide fragment bound
 CC to a HLA molecule, rather than the intact foreign antigen itself, and
 CC are particularly important in tumour rejection and in fighting viral
 CC infections. The peptides are therefore useful therapeutically to treat
 CC or prevent viral infections and cancers in mammals (especially humans)
 CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
 CC They can be administered as vaccines to elicit an immune response in
 CC individuals susceptible or otherwise at risk of viral infection or
 CC cancer, or used to treat chronic or acute conditions. They are also
 CC useful diagnostically, and can be used to induce a cytotoxic T cell
 CC response, by contacting a cytotoxic T cell with the peptide e.g. to
 CC produce CTLs ex vivo for infusion back into a patient. The
 CC polynucleotides encoding the immunogenic peptides are also useful
 CC therapeutically and for immunisation as above.
 CC
 CC Sequence 9 AA;
 SQ

Query Match 78.6%; Score 33; DB 20; Length 9;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IGHLY 7
 |||||
 Db 3 ighly 8

RESULT 7
 Y47329
 ID Y47329 standard; Peptide; 9 AA.
 XX Y47329;
 AC
 XX
 DE 01-DEC-1999 (first entry)
 XX
 DE Immunogenic peptide having a human leukocyte antigen binding motif #1940.
 XX
 KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
 KW immune response; T cell activation; major histocompatibility complex;
 KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
 KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
 KW vaccine; immunisation.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9945954-A1.
 XX
 PD 16-SEP-1999.
 XX
 PF 13-MAR-1998; 98WO-US05039.
 XX
 PR 13-MAR-1998; 98WO-US05039.
 XX
 PA (EPIM-) EPIMUNE INC.
 XX
 PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
 XX
 DR WPI; 1999-551214/46.
 XX
 PT New immunogenic peptides with HLA binding motif, useful in treatment
 PT and diagnosis of cancers and viral diseases -
 XX
 PS Claim 1; Page 103; 150pp; English.
 XX
 CC Y45390 to Y48214 represent specifically claimed immunogenic peptides
 CC having a human major histocompatibility complex (MHC) Class I (also
 CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
 CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
 CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
 CC response against the antigen from which the peptide is derived.
 CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
 CC normally induced by an antigen in the form of a peptide fragment bound
 CC to a HLA molecule, rather than the intact foreign antigen itself, and
 CC are particularly important in tumour rejection and in fighting viral
 CC infections. The peptides are therefore useful therapeutically to treat
 CC or prevent viral infections and cancers in mammals (especially humans)
 CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
 CC They can be administered as vaccines to elicit an immune response in
 CC individuals susceptible or otherwise at risk of viral infection or
 CC cancer, or used to treat chronic or acute conditions. They are also
 CC useful diagnostically, and can be used to induce a cytotoxic T cell
 CC response, by contacting a cytotoxic T cell with the peptide e.g. to
 CC produce CTLs ex vivo for infusion back into a patient. The
 CC polynucleotides encoding the immunogenic peptides are also useful
 CC therapeutically and for immunisation as above.
 CC
 CC Sequence 9 AA;
 SQ

Query Match 78.6%; Score 33; DB 20; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 IGHLYI 7
Db 2 IghlyI 7

RESULT 8
Y47950
ID Y47950 standard; Peptide; 9 AA.
XX Y47950;
XX
XX 01-DEC-1999 (first entry)
XX
XX Immunogenic peptide having a human leukocyte antigen binding motif #2561.
XX
XX Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
XX Immune response; T cell activation; major histocompatibility complex;
XX cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
XX prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
XX vaccine; immunisation.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX WO9945954-A1.
XX
XX 16-SEP-1999.
XX
XX 13-MAR-1998; 98WO-US05039.
XX
XX 13-MAR-1998; 98WO-US05039.
XX
XX (EPLM-) EPIMMUNE INC.
XX
XX Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
XX WPI, 1999-551214/46.
XX
XX New immunogenic peptides with HLA binding motif, useful in treatment
XX and diagnosis of cancers and viral diseases -
XX
XX Claim 1; Page 129; 150pp; English.
XX
XX Y45390 to Y48214 represent specifically claimed immunogenic peptides
XX having a human major histocompatibility complex (MHC) Class I (also
XX known as human leukocyte antigen (HLA)) binding motif. The immunogenic
XX peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
XX HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
XX response against the antigen from which the peptide is derived.
XX Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
XX normally induced by an antigen in the form of a peptide fragment bound
XX to a HLA molecule, rather than the intact foreign antigen itself, and
XX are particularly important in tumour rejection and in fighting viral
XX infections. The peptides are therefore useful therapeutically to treat
XX or prevent viral infections and cancers in mammals (especially humans)
XX e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
XX They can be administered as vaccines to elicit an immune response in
XX individuals susceptible or otherwise at risk of viral infection or
XX cancer, or used to treat chronic or acute conditions. They are also
XX useful diagnostically, and can be used to induce a cytotoxic T cell
XX response, by contacting a cytotoxic T cell with the peptide e.g. to
XX produce CTLs ex vivo for infusion back into a patient. The
XX polynucleotides encoding the immunogenic peptides are also useful
XX therapeutically and for immunisation as above.
XX
XX Sequence 9 AA;

Query Match 78.6%; Score 33; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 IGHLYI 7
Db 3 IghlyI 8

RESULT 9
Y46214
ID Y46214 standard; Peptide; 10 AA.
XX Y46214;
XX
XX 01-DEC-1999 (first entry)
XX
XX Immunogenic peptide having a human leukocyte antigen binding motif #825.
XX
XX Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
XX Immune response; T cell activation; major histocompatibility complex;
XX cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
XX prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
XX vaccine; immunisation.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX WO9945954-A1.
XX
XX 16-SEP-1999.
XX
XX 13-MAR-1998; 98WO-US05039.
XX
XX 13-MAR-1998; 98WO-US05039.
XX
XX (EPLM-) EPIMMUNE INC.
XX
XX Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
XX WPI, 1999-551214/46.
XX
XX New immunogenic peptides with HLA binding motif, useful in treatment
XX and diagnosis of cancers and viral diseases -
XX
XX Claim 1; Page 61; 150pp; English.
XX
XX Y45390 to Y48214 represent specifically claimed immunogenic peptides
XX having a human major histocompatibility complex (MHC) Class I (also
XX known as human leukocyte antigen (HLA)) binding motif. The immunogenic
XX peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
XX HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
XX response against the antigen from which the peptide is derived.
XX Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
XX normally induced by an antigen in the form of a peptide fragment bound
XX to a HLA molecule, rather than the intact foreign antigen itself, and
XX are particularly important in tumour rejection and in fighting viral
XX infections. The peptides are therefore useful therapeutically to treat
XX or prevent viral infections and cancers in mammals (especially humans)
XX e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
XX They can be administered as vaccines to elicit an immune response in
XX individuals susceptible or otherwise at risk of viral infection or
XX cancer, or used to treat chronic or acute conditions. They are also
XX useful diagnostically, and can be used to induce a cytotoxic T cell
XX response, by contacting a cytotoxic T cell with the peptide e.g. to
XX produce CTLs ex vivo for infusion back into a patient. The
XX polynucleotides encoding the immunogenic peptides are also useful
XX therapeutically and for immunisation as above.
XX
XX Sequence 10 AA;

Query Match 78.6%; Score 33; DB 20; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 IGHLYI 7
| | | | |
Db 4 IghlyI 9

RESULT 10
ID Y46333 standard; Peptide: 10 AA.
XX Y46333;
AC Y46333;
DT 01-DEC-1999 (first entry)
XX
XX Immunogenic peptide having a human leukocyte antigen binding motif #944.
DE
XX Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW Immune response; T cell activation; major histocompatibility complex;
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW vaccine; immunisation.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX MO9945954-A1.
PN 16-SEP-1999.
PD
XX
XX 13-MAR-1998; 98WO-US05039.
PF
XX 13-MAR-1998; 98WO-US05039.
PR
XX (EPIM-) EPIMUNE INC.
PA Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
PI WPI: 1999-551214/46.
DR
XX New immunogenic peptides with HLA binding motif, useful in treatment
PT and diagnosis of cancers and viral diseases -
PS
XX Claim 1; Page 67; 150pp; English.
XX
XX Y45390 to Y48214 represent specifically claimed immunogenic peptides
CC having a human major histocompatibility complex (MHC) Class I (also
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC response against the antigen from which the peptide is derived.
CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC normally induced by an antigen in the form of a peptide fragment bound
CC to a HLA molecule, rather than the intact foreign antigen itself, and
CC are particularly important in tumour rejection and in fighting viral
CC infections. The peptides are therefore useful therapeutically to treat
CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC They can be administered as vaccines to elicit an immune response in
CC individuals susceptible or otherwise at risk of viral infection or
CC cancer, or used to treat chronic or acute conditions. They are also
CC useful diagnostically, and can be used to induce a cytotoxic T cell
CC response, by contacting a cytotoxic T cell with the peptide e.g. to
CC produce CTLs ex vivo for infusion back into a patient. The
CC polynucleotides encoding the immunogenic peptides are also useful
CC therapeutically and for immunisation as above.
XX
XX
SQ Sequence 10 AA;

Query Match 78.6%; Score 33; DB 20; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 IGHLYI 7
| | | | |
Db 3 IghlyI 8

RESULT 11
ID Y47356 standard; Peptide: 10 AA.
XX Y47356;
AC Y47356;
DT 01-DEC-1999 (first entry)
XX
XX Immunogenic peptide having a human leukocyte antigen binding motif #1967.
DE
XX Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW Immune response; T cell activation; major histocompatibility complex;
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW vaccine; immunisation.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX MO9945954-A1.
PN 16-SEP-1999.
PD
XX
XX 13-MAR-1998; 98WO-US05039.
PF
XX 13-MAR-1998; 98WO-US05039.
PR
XX (EPIM-) EPIMUNE INC.
PA Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
PI WPI: 1999-551214/46.
DR
XX New immunogenic peptides with HLA binding motif, useful in treatment
PT and diagnosis of cancers and viral diseases -
PS
XX Claim 1; Page 105; 150pp; English.
XX
XX Y45390 to Y48214 represent specifically claimed immunogenic peptides
CC having a human major histocompatibility complex (MHC) Class I (also
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC response against the antigen from which the peptide is derived.
CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC normally induced by an antigen in the form of a peptide fragment bound
CC to a HLA molecule, rather than the intact foreign antigen itself, and
CC are particularly important in tumour rejection and in fighting viral
CC infections. The peptides are therefore useful therapeutically to treat
CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC They can be administered as vaccines to elicit an immune response in
CC individuals susceptible or otherwise at risk of viral infection or
CC cancer, or used to treat chronic or acute conditions. They are also
CC useful diagnostically, and can be used to induce a cytotoxic T cell
CC response, by contacting a cytotoxic T cell with the peptide e.g. to
CC produce CTLs ex vivo for infusion back into a patient. The
CC polynucleotides encoding the immunogenic peptides are also useful
CC therapeutically and for immunisation as above.
XX
XX
SQ Sequence 10 AA;

Query Match 78.6%; Score 33; DB 20; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 IGHLYI 7

Db 5 ighly1 10
|||||
RESULT 12
Y47357 Y47357 standard; Peptide; 10 AA.
XX
AC Y47357;
XX
DT 01-DEC-1999 (first entry)
XX
DE Immunogenic peptide having a human leukocyte antigen binding motif #1968.
XX
KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW Immune response; T cell activation; major histocompatibility complex;
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW vaccine; immunisation.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN MO9945954-A1.
XX
PD 16-SEP-1999.
XX
PF 13-MAR-1998; 98WO-US05039.
XX
PR 13-MAR-1998; 98WO-US05039.
XX
PA (EPIM-) EPIMUNE INC.
XX
PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
XX
DR WPI; 1999-551214/46.
XX
PT New immunogenic peptides with HLA binding motif, useful in treatment
XX
PS and diagnosis of cancers and viral diseases -
XX
PS Claim 1; Page 105; 150pp; English.
XX
CC Y45390 to Y48214 represent specifically claimed immunogenic peptides
CC having a human major histocompatibility complex (MHC) Class I (also
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC response against the antigen from which the peptide is derived.
CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC normally induced by an antigen in the form of a peptide fragment bound
CC to a HLA molecule, rather than the intact foreign antigen itself, and
CC are particularly important in tumour rejection and in fighting viral
CC infections. The peptides are therefore useful therapeutically to treat
CC or prevent viral infections and cancers in mammals (especially humans)
CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC They can be administered as vaccines to elicit an immune response in
CC individuals susceptible or otherwise at risk of viral infection or
CC cancer, or used to treat chronic or acute conditions. They are also
CC useful diagnostically, and can be used to induce a cytotoxic T cell
CC response, by contacting a cytotoxic T cell with the peptide e.g. to
CC produce CTLs ex vivo for infusion back into a patient. The
CC polynucleotides encoding the immunogenic peptides are also useful
CC therapeutically and for immunisation as above.
XX
SQ Sequence 10 AA;

Query Match 78.6%; Score 33; DB 20; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IGHLY1 7
|||||

Db 2 ighly1 7
|||||
RESULT 13
Y48011 Y48011 standard; Peptide; 10 AA.
XX
AC Y48011;
XX
DT 01-DEC-1999 (first entry)
XX
DE Immunogenic peptide having a human leukocyte antigen binding motif #2632.
XX
KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW Immune response; T cell activation; major histocompatibility complex;
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW vaccine; immunisation.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN MO9945954-A1.
XX
PD 16-SEP-1999.
XX
PF 13-MAR-1998; 98WO-US05039.
XX
PR 13-MAR-1998; 98WO-US05039.
XX
PA (EPIM-) EPIMUNE INC.
XX
PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
XX
DR WPI; 1999-551214/46.
XX
PT New immunogenic peptides with HLA binding motif, useful in treatment
XX
PS and diagnosis of cancers and viral diseases -
XX
PS Claim 1; Page 132; 150pp; English.
XX
CC Y45390 to Y48214 represent specifically claimed immunogenic peptides
CC having a human major histocompatibility complex (MHC) Class I (also
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC response against the antigen from which the peptide is derived.
CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC normally induced by an antigen in the form of a peptide fragment bound
CC to a HLA molecule, rather than the intact foreign antigen itself, and
CC are particularly important in tumour rejection and in fighting viral
CC infections. The peptides are therefore useful therapeutically to treat
CC or prevent viral infections and cancers in mammals (especially humans)
CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC They can be administered as vaccines to elicit an immune response in
CC individuals susceptible or otherwise at risk of viral infection or
CC cancer, or used to treat chronic or acute conditions. They are also
CC useful diagnostically, and can be used to induce a cytotoxic T cell
CC response, by contacting a cytotoxic T cell with the peptide e.g. to
CC produce CTLs ex vivo for infusion back into a patient. The
CC polynucleotides encoding the immunogenic peptides are also useful
CC therapeutically and for immunisation as above.
XX
SQ Sequence 10 AA;

Query Match 78.6%; Score 33; DB 20; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IGHLY1 7
|||||
Db 3 ighly1 8

```

RESULT 14
B08833
ID B08833 standard; peptide; 15 AA.
XX
XX B08833;
XX
XX 02-JAN-2001 (first entry)
XX
XX Amino acid sequence of a MAGE-3 derived immunogenic peptide.
DE
XX
XX Melanoma associated antigen-3; MAGE-3; immunogenic peptide;
XX major histocompatibility complex; MHC; class II molecule; tumour;
XX CD4+ T cell proliferation; melanoma cell; immune response; vaccine.
OS
XX Homo sapiens.
XX
XX W0200052045-A2.
XX
XX 08-SEP-2000.
XX
XX 23-FEB-2000; 2000WO-EP01458.
XX
XX 26-FEB-1999; 99IT-M100396.
XX
XX (SANR-) FOND CENT SAN RAFFAELE DEL MONTE TABOR.
XX
XX Protiti MP, Dellabona P;
XX
XX WPI: 2000-601867/57.
XX
XX Novel melanoma associated antigen (MAGE)-3 derived immunogenic peptides
XX useful as vaccine for inducing immune response against tumour
XX
XX Claim 1; Page 18; 27pp; English.
XX
XX B08833-37 represent melanoma associated antigen (MAGE)-3 derived
XX immunogenic peptides. The peptides are capable of binding to major
XX histocompatibility complex (MHC) class II molecules. Stimulation with
XX the peptides induces proliferation of CD4+ T cells, and of their
XX cytolytic activity. CD4+ T cells exposed to the peptides were able to
XX cause lysis of melanoma cells expressing MAGE-3 and HLA-DR molecules.
XX The peptides are useful for inducing an immune response against tumour
XX cells expressing a MAGE-3 antigen. They are also useful for the
XX preparation of anti-tumour medicament for use as a vaccine.
XX
XX Sequence 15 AA:
SQ
Query Match 78.6%; Score 33; DB 21; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 IGHLYI 7
DB 2 IGHLYI 7
RESULT 15
Y98994
ID Y98994 standard; Peptide; 15 AA.
XX
XX Y98994;
XX
XX 07-AUG-2000 (first entry)
XX
XX HLA class II binding antigen epitope peptide #183.
XX
XX Human leucocyte antigen; HLA class II; antigen epitope; pharmaceutical;
XX immune response; chronic viral disease; cancer; autoimmune disease;
XX rheumatoid arthritis; multiple sclerosis; myasthenia gravis; AIDS;
XX allograft rejection; allergy; Lyme disease; hepatitis; prostate cancer;

```

```

KW glomerulonephritis; food hypersensitivity; malaria.
XX
XX Unidentified.
OS
XX
XX W09961916-A1.
XX
XX 02-DEC-1999.
XX
XX 28-MAY-1999; 99WO-US12066.
XX
XX 29-MAY-1998; 98US-0087192.
XX
XX (EPIIM-) EPIIMONE INC.
XX
XX Sette A, Southwood S, Sidney J;
XX
XX WPI: 2000-097143/08.
XX
XX New compositions containing immunogenic peptide epitopes for various
XX HLA class II DR molecules useful for inducing helper T cell response
XX
XX Claim 1; Page 43; 60pp; English.
XX
XX The present invention relates to a new pharmaceutical composition
XX comprising a unit dose form of a peptide, or analogue, comprising an
XX epitope selected from those represented by peptides Y9812-Y9939, which
XX are derived from various antigens for various human leucocyte antigen
XX class DR molecules, representative of the world wide population. The
XX peptide/analogue binds to an HLA class II molecule at an IC-50 of less
XX than or equal to 1,000 nM. The pharmaceutical can be used to induce a
XX helper T cell response. The pharmaceutical focuses the immune response
XX towards selected determinants and could therefore be used in cases of
XX chronic viral diseases and cancer. Examples of diseases that can be
XX treated using the peptide containing pharmaceutical include autoimmune
XX diseases (rheumatoid arthritis, multiple sclerosis, and myasthenia
XX gravis), allograft rejection, allergies, Lyme disease, hepatitis,
XX post-streptococcal endocarditis or glomerulonephritis and food
XX hypersensitivities. The peptide epitopes can be used to enhance immune
XX responses against other immunogens administered with the peptides.
XX Diseases which can be treated using immunogenic mixtures include prostate
XX cancer, hepatitis B, hepatitis C, AIDS, renal carcinoma, cervical
XX carcinoma, lymphoma, and condyloma acuminatum. The peptides may also be
XX used to make monoclonal antibodies useful as potential diagnostic or
XX therapeutic agents. The peptides may also be useful as diagnostic
XX reagents, for example, to determine the susceptibility of an individual
XX to a treatment regimen. Also, the peptides may be used to predict which
XX individuals will be at substantial risk of developing chronic infection.
XX The selection of epitope based vaccines particularly towards conserved
XX development of epitope based vaccines particularly towards conserved
XX epitopes of pathogens which are characterized by high sequence
XX variability such as HIV, HCV and Malaria.
XX
XX Sequence 15 AA:
SQ
Query Match 78.6%; Score 33; DB 21; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 IGHLYI 7
DB 9 IGHLYI 14
Search completed: June 20, 2001, 14:04:02
Job time: 49 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2001, 14:04:03 ; Search time 16.63 Seconds

(without alignments)
18.539 Million cell updates/sec

Title: US-09-692-401-4

Perfect score: 46
Sequence: 1 VRIGLYTL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_39:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	314	1	MAGC_HUMAN
2	35	76.1	613	1	GCEP_CHEN
3	34	73.9	692	1	GYRB_BARBA
4	34	73.9	728	1	HGF_BARBA
5	34	73.9	728	1	HGF_MOUSE
6	34	73.9	728	1	HGF_RAT
7	33	71.7	314	1	MAG2_HUMAN
8	33	71.7	314	1	MAG3_HUMAN
9	33	71.7	619	1	GCCR_KENIA
10	32	69.6	256	1	TIFA_PHAVU
11	32	69.6	368	1	SNCL_HUMAN
12	32	69.6	415	1	ROF_HUMAN
13	32	69.6	421	1	CDS1_ARATH
14	32	69.6	449	1	ROH1_HUMAN
15	32	69.6	449	1	ROH2_HUMAN
16	32	69.6	510	1	C312_DROME
17	32	69.6	2241	1	TECU_RCMVA
18	31	67.4	358	1	LEO3_THIFE
19	31	67.4	424	1	CDS1_SOLTU
20	31	67.4	425	1	SYH_METTH
21	31	67.4	484	1	MURC_BUCAT
22	31	67.4	576	1	SYPC_YEAST
23	31	67.4	599	1	PARE_BORBU
24	31	67.4	933	1	ATID_HUMAN
25	31	67.4	1374	1	RPOB_RICPR
26	31	67.4	1376	1	RPOB_RICPR
27	30	65.2	206	1	UPP_MYCGE
28	30	65.2	206	1	UPP_MYCPN
29	30	65.2	288	1	Y077_HAEIN
30	30	65.2	303	1	POOB_ACTICA
31	30	65.2	314	1	MAG6_HUMAN
32	30	65.2	403	1	TGF_ARCFU
33	30	65.2	416	1	SYH_METJA

34	30	65.2	459	1	NEK3_HUMAN	P51956	homo sapien
35	30	65.2	510	1	G3Y3_SORYN	081973	glycine max
36	30	65.2	551	1	CC14_YEAST	000684	saccharomyc
37	30	65.2	683	1	EFGL_THIMA	Q9x1y4	thermopoga
38	30	65.2	883	1	RPOL_BPT7	P00573	bacterioph
39	30	65.2	906	1	RPOL_BPK11	P18147	bacterioph
40	30	65.2	3898	1	POLG_BVDYN	001499	bovine vira
41	30	65.2	3988	1	POLG_BVDYN	P19711	bovine vira
42	29	63.0	76	1	RPOB_ARCFU	029135	archaeoglob
43	29	63.0	167	1	ISPA_CHLTR	084413	chlamydia t
44	29	63.0	193	1	NADD_BORBU	051723	borrelia bu
45	29	63.0	230	1	YME1_CVHOC	Q01455	human coron

ALIGNMENTS

```

RESULT 1
MAGC_HUMAN STANDARD: PRT; 314 AA.
AC P43365;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE MELANOMA-ASSOCIATED ANTIGEN 12 (MAGE-12 ANTIGEN) (MAGE12F).
GN MAGE12 OR MAGE12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94102805; PubMed=8276455;
RA de Smet C., Turquin C., van der Bruggen P., de Plaen E., Brasseur F.,
RA Boon T.;
RT "Sequence and expression pattern of the human MAGE2 gene.";
RL Immunogenetics 39:121-129(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Skin;
RX MEDLINE=94311935; PubMed=8037761;
RA Ding M., Beck R.U., Keller C.U., Fenton R.G.;
RT "Cloning and analysis of MAGE-1-related genes.";
RL Biochem. Biophys. Res. Commun. 202:549-555(1994).
CC -1- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN TUMOR
CC TRANSFORMATION OR ASPECTS OF TUMOR PROGRESSION.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES.
CC -1- SIMILARITY: BELONGS TO THE MAGE FAMILY. STRONG SIMILARITY (90%)
CC WITH MAGE-2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L18877; AAA19023.1; -.
DR MIR: 300177; -.
DR InterPro: IPR002190; -.
DR Pfam: PF01454; MAGE; 1.
KW Antigen; Multigene family; Tumor antigen.
FT DOMAIN 40 43 POLY-SER.
FT CONFICT 10 10 C -> S (IN REF. 2).
FT CONFICT 187 187 A -> D (IN REF. 2).
FT CONFICT 300 300 P -> S (IN REF. 2).
SQ SEQUENCE 314 AA; 34802 MW; 3F0787CECD8816A5 CRC64;

```

Query Match 100.0%; Score 46; DB 1; Length 314;
 Best Local Similarity 100.0%; Pred. No. 0.056;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRIGHLYI 9
 |||||
 DB 170 VRIGHLYI 178

RESULT 2
 GCPE.CHLPN STANDARD; PRT; 613 AA.

AC 0928H0; Q93Q95; 401, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GCPE PROTEIN HOMOLOG.
 GN GCPE OR CPN0373 OR CP0383.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiales; Chlamydiae; Chlamydia.
 OX NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CWL029;
 RX MEDLINE=99206606; PubMed=10192388;
 RA Kalman S., Mitchell W., Marathe R., Lamme C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis";
 RL Nat. Genet. 21:385-389(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=201550255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,
 RA Linner K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Eissen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 RT pneumoniae AR39";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shitai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shida T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CWL029 from USA";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -1- SIMILARITY: BELONGS TO THE GCPE FAMILY
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE001621; AAD18517.1; -
 CC EMBL: AE002200; AAF38230.1; -
 CC EMBL: AP002546; BAA98581.1; -
 CC TIGR: CP0383; -
 CC SEQUENCE 613 AA; 68308 MW; 40B1F831E02DC6D4 CRC64;

Query Match 76.1%; Score 35; DB 1; Length 613;
 Best Local Similarity 87.5%; Pred. No. 17;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VRIGHLYI 8
 |||||

DB 18 VRIGHLYI 25

RESULT 3
 GYRB_BARBA STANDARD; PRT; 692 AA.

AC P94281;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE DNA GYRASE SUBUNIT B (EC 5.99.1.3).
 GN GYRB.
 OS Bartonella bacilliformis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Bartonellaceae; Bartonella.
 OX NCBI_TaxID=774;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KC583 / ATCC 35685;
 RX MEDLINE=99013570; PubMed=9797224;
 RA Battisti J.M., Smithe L.S., Samuels D.S., Minnick M.F.;
 RT "Mutations in Bartonella bacilliformis gyrb confer resistance to
 RT coumermycin A1";
 RL Antimicrob. Agents Chemother. 42:2906-2913(1998).
 CC -1- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-
 CC STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE
 CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
 CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
 CC -1- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING
 CC OF DOUBLE-STRANDED DNA.
 CC -1- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
 CC BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
 CC ENZYME FORMS AN A2B2 TETRAMER.
 CC -1- SIMILARITY: BELONGS TO TYPE II TOPOISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U82225; AAC71079.1; -
 CC InterPro: IPR001241; -
 CC DR InterPro: IPR002288; -
 CC DR Pfam: PF00986; DNA_gyraseB_C; 1.
 CC DR Pfam: PF00204; DNA_topoisomI; 1.
 CC DR PRINTS: PR00418; TP12FAMILY.
 CC DR PROSITE: PS00177; TOPOISOMERASE_II; 1.
 CC KW Topoisomerase; Isomerase; ATP-binding; Antibiotic resistance.
 CC FT VARIANT 124 124 G->S (IN COUMERMYCIN A1 MUTANT).
 CC FT VARIANT 184 184 R->Q (IN COUMERMYCIN A1 MUTANT).
 CC FT VARIANT 214 214 T->A (IN COUMERMYCIN A1 MUTANT).
 CC FT VARIANT 214 214 T->I (IN COUMERMYCIN A1 MUTANT).
 CC SEQUENCE 692 AA; 77445 MW; F54948EEF342281F CRC64;

Query Match 73.9%; Score 34; DB 1; Length 692;
 Best Local Similarity 62.5%; Pred. No. 30;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VRIGHLYI 8
 |||||
 DB 576 IRAGHLYI 583

RESULT 4
 HGF_HUMAN STANDARD; PRT; 728 AA.

AC P14210;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HEPATOCYTE GROWTH FACTOR PRECURSOR (SCATTER FACTOR) (SF)
 DE (HEPATOPOIETIN A).
 GN HGF OR HPTA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91340155; PubMed-1831432;
 RA Seki T., Hagiya M., Shimonishi M., Nakamura T., Shimizu S.;
 RT "Organization of the human hepatocyte growth factor-encoding gene";
 RL Gene 102:213-219(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Placenta;
 RX MEDLINE-89392017; PubMed-2528952;
 RA Miyazawa K., Tsubouchi H., Naka D., Takahashi K., Okigaki M.,
 RA Arakaki N., Nakayama H., Hirono S., Sakiyama O., Takahashi K.,
 RA Gonda E., Daikuhara Y., Kitamura N.;
 RT "Molecular cloning and sequence analysis of cDNA for human hepatocyte
 growth factor.";
 RL Biochem. Biophys. Res. Commun. 163:967-973(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Leukocyte;
 RX MEDLINE-91025062; PubMed-2145836;
 RA Seki T., Ihara I., Sugimura A., Shimonishi M., Nishizawa T.,
 RA Asami O., Hagiya M., Nakamura T., Shimizu S.;
 RT "Isolation and expression of cDNA for different forms of hepatocyte
 growth factor from human leukocyte.";
 RL Biochem. Biophys. Res. Commun. 172:321-327(1990).
 RN [4]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 55-73 AND 495-520.
 RC TISSUE-Liver;
 RX MEDLINE-90066676; PubMed-2531289;
 RA Nakamura T., Nishizawa T., Hagiya M., Seki T., Shimonishi M.,
 RA Sugimura A., Tashiro K., Shimizu S.;
 RT "Molecular cloning and expression of human hepatocyte growth factor.";
 RL Nature 342:440-443(1989).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Embryonic fibroblast;
 RX MEDLINE-91334393; PubMed-1831266;
 RA Weidner K.M., Arakaki N., Hartmann G., Vandeckerckhove J., Weingart S.,
 RA Rieder H., Forstsch C., Tsubouchi H., Hishida T., Daikuhara Y.,
 RA Birchemer M.;
 RT "Evidence for the identity of human scatter factor and human
 hepatocyte growth factor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:7001-7005(1991).
 RN [6]
 RP SIGNAL SEQUENCE CLEAVAGE SITE.
 RX MEDLINE-91207365; PubMed-1826837;
 RA Yoshiyama Y., Arakaki N., Naka D., Takahashi K., Hirono S., Kondo J.,
 RA Nakayama H., Gonda E., Kitamura N., Tsubouchi H., Ishii T.,
 RA Hishida T., Daikuhara Y.;
 RT "Identification of the N-terminal residue of the heavy chain of both
 native and recombinant human hepatocyte growth factor.";
 RL Biochem. Biophys. Res. Commun. 175:660-667(1991).
 RN [7]
 RP CARBOHYDRATE-LINKAGE SITE 476.
 RX MEDLINE-93129192; PubMed-1482348;
 RA Shimizu N., Hara H., Sogabe T., Sakai H., Ihara I., Inoue H.,
 RA Nakamura T., Shimizu S.;
 RT "Hepatocyte growth factor is linked by O-glycosylated oligosaccharide
 on the alpha chain.";
 RL Biochem. Biophys. Res. Commun. 189:1329-1335(1992).
 RN [8]
 RP MUTAGENESIS.
 RX MEDLINE-92331602; PubMed-1321034;
 RA Lokker N.A., Mark M.R., Luis E.A., Bennett G.L., Robbins K.A.,
 RA Baker J.B., Godowski P.J.;

RT "Structure-function analysis of hepatocyte growth factor:
 RT Identification of variants that lack mitogenic activity yet retain
 RT high affinity receptor binding.";
 RL EMBO J. 11:2503-2510(1992).
 RN [9]
 RP STRUCTURE BY NMR OF 31-127.
 RX MEDLINE-98154323; PubMed-9493272;
 RA Zhou H., Mazzulla M.J., Kaufman J.D., Stahl S.J., Wingfield P.T.,
 RA Rubin J.S., Bottaro D.P., Byrd R.A.;
 RT "The solution structure of the N-terminal domain of hepatocyte growth
 RT factor reveals a potential heparin-binding site.";
 RL Structure 6:109-116(1998).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 35-210.
 RX MEDLINE-99036958; PubMed-9817840;
 RA Ullsch M., Lokker N.A., Godowski P.J., de Vos A.M.;
 RT "Crystal structure of the NK1 fragment of human hepatocyte growth
 RT factor at 2.0-A resolution.";
 RL Structure 6:1383-1393(1998).
 CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
 CC HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPHIC FACTOR, AND ACTS
 CC AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
 CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.
 CC -1- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
 CC DISULFIDE BOND.
 CC -1- SIMILARITY: CONTAINS 4 KRINGLE REGIONS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1: ALSO KNOWN AS THE
 CC TRIPSLIN FAMILY, PLASMINOGEN SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D90334; BAA14348.1; JOINED.
 DR EMBL; D90318; BAA14348.1; JOINED.
 DR EMBL; D90319; BAA14348.1; JOINED.
 DR EMBL; D90320; BAA14348.1; JOINED.
 DR EMBL; D90322; BAA14348.1; JOINED.
 DR EMBL; D90323; BAA14348.1; JOINED.
 DR EMBL; D90324; BAA14348.1; JOINED.
 DR EMBL; D90325; BAA14348.1; JOINED.
 DR EMBL; D90326; BAA14348.1; JOINED.
 DR EMBL; D90327; BAA14348.1; JOINED.
 DR EMBL; D90328; BAA14348.1; JOINED.
 DR EMBL; D90329; BAA14348.1; JOINED.
 DR EMBL; D90330; BAA14348.1; JOINED.
 DR EMBL; D90331; BAA14348.1; JOINED.
 DR EMBL; D90332; BAA14348.1; JOINED.
 DR EMBL; D90333; BAA14348.1; JOINED.
 DR EMBL; M29145; AAA52650.1; -.
 DR EMBL; M60718; AAA52648.1; -.
 DR EMBL; X16323; CAA34387.1; -.
 DR EMBL; M73259; AAA64239.1; -.
 DR EMBL; M73240; AAA64237.1; -.
 DR PIR; JH0579; JH0579.
 DR PIR; S06794; S06794.
 DR PDB; 2HGF; 24-JUN-98.
 DR PDB; 1BHT; 18-NOV-98.
 DR MEROPS; S01.976; -.
 DR GlycositeDB; P14210; -.
 DR MIM; 142409; -.
 DR InterPro; IPR000001; -.
 DR InterPro; IPR001254; -.
 DR InterPro; IPR001314; -.
 DR InterPro; IPR003014; -.
 DR Pfam; PF00024; PAN; 1.
 DR Pfam; PF00051; kringle; 4.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00018; KRINGLE.

FT CARBOHYD 656 656 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 163 167 MISSING (IN SHORT ISOFORM).
 FT CONFLICT 344 344 N -> K (IN REF. 2).
 FT CONFLICT 479 479 V -> L (IN REF. 2).
 FT CONFLICT 564 564 R -> H (IN REF. 3).
 FT CONFLICT 564 564
 SO SEQUENCE 728 AA; 82944 MW; A0381FCA954328 CRC64;

Query Match 73.9%; Score 34; DB 1; Length 728;
 Best Local Similarity 55.6%; Pred. No. 31;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Yy 1 VRIGHLYIL 9
 Db 632 LRVAHLXIM 640

RESULT 6
 HGF_RAT STANDARD; PRT; 728 AA.
 AC P17945;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE HEPATOCYTE GROWTH FACTOR PRECURSOR (SCATTER FACTOR) (SF)
 DE (HEPATOPOETIN A).
 GN HGF.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=MISTAR; TISSUE=Liver;
 RX MEDLINE=90222197; PubMed=2139229;
 RA Toshio K., Hagiya M., Nishizawa T., Seki T., Shimonishi M.,
 RA Shimizu S., Nakamura T.;
 RT "Deduced primary structure of rat hepatocyte growth factor and
 RT expression of the mRNA in rat tissues";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3200-3204(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MISTAR; TISSUE=Liver;
 RX MEDLINE=91031482; PubMed=2146117;
 RA Okajima A., Miyazawa K., Kitamura N.;
 RT "Primary structure of rat hepatocyte growth factor and induction of
 RT its mRNA during liver regeneration following hepatic injury";
 RL Eur. J. Biochem. 193:375-381(1990)
 CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
 CC HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPHIC FACTOR, AND ACTS
 CC AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
 CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.
 CC -1- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
 CC DISULFIDE BOND.
 CC -1- SIMILARITY: CONTAINS 4 KRINGLE REGIONS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY PLASMINOGEN SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: D90102; BA014133.1; -;
 CC EMBL: X54400; CAA38266.1; -;
 CC PIR: S13211; S13211.
 CC PIR: A35644; A35644.
 CC HSSP: P14210; 2HGF.
 CC MEROPS: S01.978; -;
 CC InterPro: IPR000001; -;

DR InterPro: IPR001254; -;
 DR InterPro: IPR001314; -;
 DR InterPro: IPR003014; -;
 DR Pfam: PF00024; PAN; 1.
 DR Pfam: PF00051; kringle; 4.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PRO0018; KRINGLE.
 DR PRINTS: PRO0722; CHIMOTRYPSIN.
 DR PROSITE: PS00021; KRINGLE_1; 4.
 DR PROSITE: PS00070; KRINGLE_2; 4.
 KW Growth factor; Kringle; Glycoprotein; Serine protease homolog;
 KW Signal.

FT SIGNAL 1 32 BY SIMILARITY.
 FT CHAIN 33 485 HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.
 FT CHAIN 496 728 HEPATOCYTE GROWTH FACTOR BETA CHAIN.
 FT MOD_RES 33 33 PYRROLIDONE CARBOXYLIC ACID
 (BY SIMILARITY).
 FT DOMAIN 33 128 PAP.
 FT DOMAIN 129 207 KRINGLE 1.
 FT DOMAIN 212 289 KRINGLE 2.
 FT DOMAIN 306 384 KRINGLE 3.
 FT DOMAIN 392 470 KRINGLE 4.
 FT DOMAIN 496 728 SERINE PROTEASE-LIKE.
 FT DISULFID 71 97 BY SIMILARITY.
 FT DISULFID 75 85 BY SIMILARITY.
 FT DISULFID 488 607 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 656 656 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 728 AA; 82905 MW; 3E0BF196ADCEDEF CRC64;

Query Match 73.9%; Score 34; DB 1; Length 728;
 Best Local Similarity 55.6%; Pred. No. 31;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Yy 1 VRIGHLYIL 9
 Db 632 LRVAHLXIM 640

RESULT 7
 MAG2_HUMAN STANDARD; PRT; 314 AA.
 AC P43356;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE MELANOMA-ASSOCIATED ANTIGEN 2 (MAGE-2 ANTIGEN).
 DE MAGE2 OR MAGE2.
 GN Homo sapiens (Human).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=94102805; PubMed=8276455;
 RA de Smet C., Lurquin C., van der Bruggen P., de Plaen E., Brasseur F.,
 RA Boon T.;
 RT "Sequence and expression pattern of the human MAGE2 gene";
 RL Immunogenetics 39:121-129(1994).
 RN [2]
 RP MUTAGENESIS.
 RC TISSUE=Blood;
 RX MEDLINE=94157413; PubMed=8113684;
 RA Gauglier B., van den Eynde B., van der Bruggen P., Romero P.,
 RA Gafioro J.J., de Plaen E., Lethe B., Brasseur F., Boon T.;
 RT "Human gene MAGE-3 codes for an antigen recognized on a melanoma by
 RT autologous cytolytic T lymphocytes";
 RL J. Exp. Med. 179:921-930(1994).
 CC -1- FUNCTION: NOT KNOWN. THOUGH MAY PLAY A ROLE IN EMBRYONAL
 CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR

```

CC PROGRESSION. ANTIGEN RECOGNIZED ON A MELANOMA BY AUTOLOGOUS
CC CYTOLYTIC T LYMPHOCYTES.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES.
CC -1- SIMILARITY: BELONGS TO THE MAGE FAMILY. STRONG SIMILARITY (90%)
CC WITH MAGE-12.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: L18920; AAA17729.1; -.
CC DR MIM: 300173; -.
CC DR InterPro: IPR002190; -.
CC DR Pfam: PF01454; MAGE; 1.
CC KW Antigen; Multigene family; Tumor antigen.
CC FT DOMAIN 40 43 POLY-SER.
CC FT MUTAGEN 170 170 V->D: IMPROVES ABILITY TO BIND TO HLA-A1.
CC SQ SEQUENCE 314 AA; 35055 MW; 844F16335A2BCE7 CRC64;

Query Match 71.7%; Score 33; DB 1; Length 314;
Best Local Similarity 77.8%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 VRIGHLTYL 9
Db 170 VPISHLYL 178

RESULT 8
MAG3_HUMAN STANDARD; PRT; 314 AA.
AC P43357;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE MELANOMA-ASSOCIATED ANTIGEN 3 (MAGE-3 ANTIGEN) (ANTIGEN M22-D).
GN MAGE3 OR MAGE3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. AND MUTAGENESIS.
RC TISSUE=Blood;
RX MEDLINE=94157413; PubMed=8113684;
RA Gafaric B., van den Eynde B., van der Bruggen P., Romero P.,
RA Gafaric J.J., de Plaen E., Lethe B., Brassens F., Boon T.;
RT "Human gene MAGE-3 codes for an antigen recognized on a melanoma by
RT autologous cytolytic T lymphocytes.";
RL J. Exp. Med. 179:921-930(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=94311935; PubMed=8037761;
RA Ding M., Beck R.J., Keller C.J., Fenton R.G.;
RT "Cloning and analysis of MAGE-1-related genes.";
RT Biochem. Biophys. Res. Commun. 202:549-555(1994).
CC -1- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
CC PROGRESSION. ANTIGEN RECOGNIZED ON A MELANOMA BY AUTOLOGOUS
CC CYTOLYTIC T LYMPHOCYTES.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES AND PLACENTA. NEVER EXPRESSED IN KIDNEY TUMORS,

```

```

CC LEUKEMIAS AND LYMPHOMAS.
CC -1- SIMILARITY: BELONGS TO THE MAGE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: U03735; AAA17446.1; -.
CC DR MIM: 300174; -.
CC DR InterPro: IPR002190; -.
CC DR Pfam: PF01454; MAGE; 1.
CC KW Antigen; Multigene family; Tumor antigen.
CC FT DOMAIN 40 43 POLY-SER.
CC FT MUTAGEN 170 170 D->A: ABOLISHES HLA-A1 BINDING.
CC FT MUTAGEN 176 176 Y->A: ABOLISHES HLA-A1 BINDING.
CC SQ SEQUENCE 314 AA; 34747 MW; 3F5EB13D1C946A1 CRC64;

Query Match 71.7%; Score 33; DB 1; Length 314;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 IGHLYT 8
Db 172 IGHLYT 177

RESULT 9
GCKR_XENLA STANDARD; PRT; 619 AA.
AC O91754;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GLUCOKINASE REGULATORY PROTEIN (GLUCOKINASE REGULATOR).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodina; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95010134; PubMed=7925465;
RA Velga-Da-Cunha M., Delheux M., Watelet N., van Schaftingen E.;
RT "Cloning and expression of a xenopus liver cDNA encoding a fructose-
RT phosphate-insensitive regulatory protein of glucokinase.";
RL Eur. J. Biochem. 225:43-51(1994).
CC -1- FUNCTION: INHIBITS GLUCOKINASE BY FORMING AN INACTIVE COMPLEX WITH
CC THIS ENZYME.
CC -1- SIMILARITY: BELONGS TO THE SIS FAMILY. GCKR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: X80901; CAA56863.1; -.
CC DR InterPro: IPR001347; -.
CC DR InterPro: IPR001741; -.
CC DR Pfam: PF01380; SIS; 1.
CC DR PROSITE: PS01272; GCKR; 1.
CC SQ SEQUENCE 619 AA; 68738 MW; 41B72C1981D1BA52 CRC64;

Query Match 71.7%; Score 33; DB 1; Length 619;

```

Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GHLYIL 9
|||||

Db 327 GHLYIL 332

RESULT 10
ID TIPA_PRAVU STANDARD; PRT; 256 AA.
AC P23958;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE TONOPLAST INTRINSIC PROTEIN, ALPHA (ALPHA TIP).
OS Phaseolus vulgaris (kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Phaseolus.
OX NCBI_Taxid=3885;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 11-21.
RC SPRAIN-CV, GREENSLIEVES: TISSUE=Seed;
RX MEDLINE=93044491; PubMed=2152174;
RA Johnson K.D., Hoeft H.R., Chrispeels M.J.;
RT "An intrinsic tonoplast protein of protein storage vacuoles in seeds
is structurally related to a bacterial solute transporter (GIPF).";
RL Plant Cell 2:525-532(1990).
RN [2]
RP PHOSPHORYLATION BY CDPK.
RA Johnson K.D., Chrispeels M.J.;
RT "Tonoplast-bound protein kinase phosphorylates tonoplast intrinsic
protein."
RL Plant Physiol. 100:1787-1795(1992).
CC -1- FUNCTION: CHANNEL PROTEIN IN TONOPLAST. THESE PROTEINS MAY ALLOW
THE DIFFUSION OF AMINO ACIDS AND/OR PEPTIDES FROM THE TONOPLAST
INTERIOR TO THE CYTOPLASM.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: FOUND IN ALL SEED TISSUES THAT ARE ALIVE AT
SEED MATURITY, BUT NOT IN TISSUES THAT LOSE VIABILITY DURING
SEED MATURATION.
CC -1- PFM: PHOSPHORYLATED BY A TONOPLAST-BOUND CALCIUM-DEPENDENT
PROTEIN KINASE.
CC -1- SIMILARITY: BELONGS TO THE TRANSMEMBRANE CHANNEL MIP FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X62873; CAA44669.1; -
DR PIR: J01106; J01106.
DR PIR: S26742; S26742.
DR Interpro: IPR000435; -
DR Pfam: PF00230; MIP_1.
DR PRINTS: PR00783; MINTRINSCP.
DR PROSITE: PS00221; MIP_1.
KW Transport; Transmembrane; Multigene family; Seed; Phosphorylation.
FT DOMAIN 1 24 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 25 44 POTENTIAL.
FT DOMAIN 45 57 VACUOLAR (POTENTIAL).
FT TRANSSEM 58 77 POTENTIAL.
FT DOMAIN 78 102 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 103 121 POTENTIAL.
FT DOMAIN 122 143 VACUOLAR (POTENTIAL).
FT TRANSSEM 144 164 POTENTIAL.
FT DOMAIN 165 169 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 170 189 POTENTIAL.
FT DOMAIN 190 216 VACUOLAR (POTENTIAL).

FT TRANSSEM 217 239 POTENTIAL.
FT DOMAIN 240 256 CYTOPLASMIC (POTENTIAL).
FT MOD_RES 7 7 PHOSPHORYLATION (BY CDPK) (PROBABLE).
SQ SEQUENCE 256 AA; 27195 MW; 7862A0AC8E846810 CRC64;

Query Match 69.6%; Score 32; DB 1; Length 256;
Best Local Similarity 55.6%; Pred. No. 27;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VRIGHLYIL 9
|:|:|:|
Db 139 VGVGHMFIL 147

RESULT 11
ID SNCL_HUMAN STANDARD; PRT; 368 AA.
AC Q16533;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SNRNA ACTIVATING PROTEIN COMPLEX 43 KDA SUBUNIT (SNAPC 43 KDA
SUBUNIT) (PROXIMAL SEQUENCE ELEMENT-BINDING TRANSCRIPTION FACTOR GAMMA
DE SUBUNIT) (PSE-BINDING FACTOR GAMMA SUBUNIT) (PTF GAMMA SUBUNIT).
CN SNAPC1 OR SNAP43.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=95231630; PubMed=7715707;
RA Henry R.W., Sadowski C.L., Kobayashi R., Hernandez N.;
RT "A TBP-TAF complex required for transcription of human snRNA genes by
RNA polymerase II and III".
RL Nature 374:653-656(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96104548; PubMed=8524284;
RA Yoon J.B., Roeder R.G.;
RT "Cloning of two proximal sequence element-binding transcription factor
subunits (gamma and delta) that are required for transcription of
small nuclear RNA genes by RNA polymerases II and III and interact
with the TATA-binding protein.";
RL Mol. Cell. Biol. 16:1-9(1996).
CC -1- FUNCTION: COMPLEX REQUIRED FOR THE TRANSCRIPTION OF BOTH RNA
POLYMERASE II AND III SMALL-NUCLEAR RNA GENES. BINDS TO THE
PROXIMAL SEQUENCE ELEMENT (PSE), A NON-TATA-BOX BASAL PROMOTER
ELEMENT COMMON TO THESE 2 TYPES OF GENES.
CC -1- SUBUNIT: COMPOSED OF 5 SUBUNITS: SNAP190; SNAP50; SNAP45; SNAP43
AND SNAP19.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z47542; CA87590.1; -
DR EMBL: U44754; AAC50358.1; -
DR MIM: 600591; -
DR TRANSFAC: T01680; -
KW Transcription regulation.
FT DOMAIN 347 352 POLY-GLU.
SQ SEQUENCE 368 AA; 42994 MW; 324E89C8FB540C32 CRC64;

Query Match 69.6%; Score 32; DB 1; Length 368;
Best Local Similarity 55.6%; Pred. No. 39;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 WRIGHTYL 9
:1:11:1
DB 74 INVAGLYL 82

RESULT 12
ID ROF_HUMAN STANDARD; PRT; 415 AA.
AC P52597.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1998 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN F (HNRNP F).
GN HNRPF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 51-60; 77-82 AND 127-135.
RX MEDLINE=94203790; PubMed=7512260;
RA Matunis M.J., Xing J., Dreyfuss G.;
RT "The hnrnp F protein: unique primary structure, nucleic acid-binding
RT properties, and subcellular localization.";
RL Nucleic Acids Res. 22:1059-1067(1994).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96081943; PubMed=7499401;
RA Honore B., Rasmussen H.H., Vorum H., Deigaard K., Liu X.,
RA Gromov P., Madsen P., Gesser B., Tommerup N., Celis J.E.;
RT "Heterogeneous nuclear ribonucleoproteins H, H', and F are members of
RT a ubiquitously expressed subfamily of related but distinct proteins
RT encoded by genes mapping to different chromosomes.";
RL J. Biol. Chem. 270:28780-28789(1995).
CC -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE HETEROGENEOUS NUCLEAR
CC RIBONUCLEOPROTEIN (HNRNP) COMPLEXES WHICH PROVIDE THE SUBSTRATE
CC FOR THE PROCESSING EVENTS THAT PRE-MRNAs UNDERGO BEFORE BECOMING
CC FUNCTIONAL, TRANSLATABLE MRNAs IN THE CYTOPLASM. PROBABLY BINDS G-
CC RICH SEQUENCES IN PRE-MRNAs.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC -1- TISSUE SPECIFICITY: EXPRESSED UBQUITOUSLY.
CC -1- SIMILARITY: CONTAINS 3 RNA RECOGNITION MOTIFS (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L28010; AAC37584.1; -
CC DR Aairhus/Ghent-2DPAGE; 6504; IEF.
CC DR Aairhus/Ghent-2DPAGE; 7312; IEF.
CC DR MIM; 601037; -
CC DR InterPro: IPR000504; -
CC DR Pfam: PF00076; rrm; 3.
CC DR PROSITE: PS01002; rrm; 2.
CC DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
CC DR Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat.
CC KW DOMAIN 11 ? RNA-BINDING (RRM) 1.
CC FT DOMAIN 111 188 RNA-BINDING (RRM) 2.
CC FT DOMAIN 289 366 RNA-BINDING (RRM) 3.
CC FT SEQUENCE 415 AA; 45672 MW; D14EL70631FB1F31 CRC64;

Query Match 69.6%; Score 32; DB 1; Length 415;
Best Local Similarity 85.7%; Pred. No. 44;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RIGHTYL 8
111111

DB 175 RIGHTYL 181

RESULT 13
ID CDS1_ARATH STANDARD; PRT; 421 AA.
AC 004928; O48608;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PHOSPHATIDATE CYTIDYLTRANSFERASE (EC 2.7.7.41) (CDP-DIGLYCERIDE
DE SYNTHETASE) (CDP-DIGLYCERIDE PHOSPHOTRANSFERASE) (CDP-DIACYLGlycerol
DE SYNTHASE) (CDS) (CYP:PHOSPHATIDATE CYTIDYLTRANSFERASE) (CDP-DAG
DE SYNTHASE) (CDP-DG SYNTHETASE).
GN CDS1 OR F2401.17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=CV. C24; TISSUE=flower, and Siliques;
RX MEDLINE=97239925; PubMed=9085581;
RA Kopka J., Ludwig M., Mueller-Roeber B.;
RT "Complementary DNAs encoding eukaryotic-type cytidine-5'-diphosphate-
RT diacylglycerol synthases of two plant species.";
RL Plant Physiol. 113:1997-1002(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Chauk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altati H., Bel B., Chin C., Chou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsy N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Scharif J., Southwick A.,
RA Thavert A., Tortum M., Vaynsberg K., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY BE INVOLVED IN THE SYNTHESIS OF MINOR PHOSPHOLIPIDS
CC AND IN MODULATION OF IP3-MEDIATED SIGNAL TRANSDUCTION.
CC -1- CATALYTIC ACTIVITY: CTP + PHOSPHATIDATE = PYROPHOSPHATE +
CC CDP-DIACYLGlycerol.
CC -1- PATHWAY: PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE CDS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X94306; CA63969.1; -
CC DR EMBL: AC003113; AAF70845.1; -
CC DR Mendel; 27238; Arath; 2834; 27238.
CC DR InterPro: IPR000374; -
CC DR Pfam: PF01148; Cytidylyltrans; 1.
CC DR PROSITE: PS01315; CDS; 1.
CC KW Transferrase; Nucleotidyltransferase; Phospholipid biosynthesis;
CC KW Transmembrane.
CC FT TRANSMEM 60 80 POTENTIAL.
CC FT TRANSMEM 149 169 POTENTIAL.
CC FT TRANSMEM 206 226 POTENTIAL.
CC FT TRANSMEM 246 266 POTENTIAL.
CC FT TRANSMEM 321 341 POTENTIAL.
CC FT SEQUENCE 421 AA; 48659 MW; 96D238DB4B913A3 CRC64;

Query Match 69.6%; Score 32; DB 1; Length 421;
Best Local Similarity 75.0%; Pred. No. 45;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VRIGHTX1 8
1 :|||||
Db 67 VYMGHLY1 74

RESULT 14
ROH1_HUMAN STANDARD: PRT: 449 AA.
AC P31943;
DT 01-VOL-1993 (Rel. 26, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN H (HNRNP H).
GN HNRPH1 OR HNRPH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 180-184; 193-197 AND 200-230.
RX MEDLINE=96081943; PubMed=7499401;
RA Honore B., Rasmussen H.H., Vorum H., Dejgaard K., Liu X.,
RA Gromov P., Madsen P., Gesser B., Tommerup N., Celis J.E.;
RT "Heterogeneous nuclear ribonucleoproteins H, H', and F are members of
RT a ubiquitously expressed subfamily of related but distinct proteins
RT encoded by genes mapping to different chromosomes.";
RL J. Biol. Chem. 270:28780-28789(1995).
RN [2]
RP SEQUENCE OF 127-135 AND 153-163.
RX MEDLINE=94203790; PubMed=7512260;
RA Matunis M.J., Xing J., Dreyfuss G.;
RT "The hnrnp F protein: unique primary structure, nucleic acid-binding
RT properties, and subcellular localization.";
RL Nucleic Acids Res. 22:1059-1067(1994).
RN [3]
RP SEQUENCE OF 200-230.
RX TISSUE=keratinocytes;
RA MEDLINE=93162043; PubMed=1286667;
RA Rasmussen H.H., Van Damme J., Puype M., Gesser B., Celis J.E.,
RA Vandekerckhove J.;
RT "Microsequences of 145 proteins recorded in the two-dimensional gel
RT protein database of normal human epidermal keratinocytes.";
RL Electrophoresis 13:960-969(1992).
RN [4]
RP FUNCTION: THIS PROTEIN IS A COMPONENT OF THE HETEROGENEOUS NUCLEAR
RN RIBONUCLEOPROTEIN (HNRNP) COMPLEXES WHICH PROVIDE THE SUBSTRATE
RN FOR THE PROCESSING EVENTS THAT PRE-MRNAS UNDERGO BEFORE BECOMING
RN FUNCTIONAL, TRANSLATABLE MRNAS IN THE CYTOPLASM. BINDS POLY (RG).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOPLASM.
CC -1- TISSUE SPECIFICITY: EXPRESSED UBQUITOUSLY.
CC -1- DOMAIN: EACH QUASI-RRM REPEAT BOUND POLY(RG), WHILE ONLY THE N-
CC TERMINAL RRM BOUND POLY(RC) AND POLY (RU). NONE OF THE REPEATS
CC BOUND DETECTABLE AMOUNTS OF POLY(RA).
CC -1- SIMILARITY: CONTAINS 3 RNA RECOGNITION MOTIFS (RRM).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L22009; AAA91346.1;
DR Aairnus/Genet-2DPAGE; 4410; IEF.
DR Aairnus/Genet-2DPAGE; 4429; IEF.
DR Aairnus/Genet-2DPAGE; 5416; IEF.
DR MIM: 601035;
DR InterPro: IPR000504;
DR Pfam: PF00076; rrm; 3.
DR PROSITE: PS50102; RRM; 3.
DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG
KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat.

FT DOMAIN 11 90 RNA-BINDING (RRM) 1.
FT DOMAIN 111 188 RNA-BINDING (RRM) 2.
FT DOMAIN 289 364 RNA-BINDING (RRM) 3.
FT DOMAIN 234 433 2 X 16 AA GLY-RICH APPROXIMATE REPEATS.
FT REPEAT 234 249 1-1.
FT REPEAT 418 433 1-2.
FT DOMAIN 354 392 2 X 19 AA PERFECT REPEATS.
FT REPEAT 354 372
FT REPEAT 374 392 2-1.
SQ SEQUENCE 449 AA; 49229 MW; 4ECF7A075C2526FF CRC64;

Query Match 69.6%; Score 32; DB 1; Length 449;
Best Local Similarity 85.7%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 RICHLY1 8
1 :|||||
Db 175 RICHLY1 181

RESULT 15
ROH2_HUMAN STANDARD: PRT: 449 AA.
AC P55795;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN H' (HNRNP H') (FTF-3).
GN HNRPH2 OR FTF3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95048329; PubMed=7959728;
RA Vorechovsky I., Veltrie D., Holland J., Bentley D.R., Thomas K.,
RA Zhou J.N., Notarangelo L.D., Plehani A., Fontan G., Ochs H.D.;
RT "Isolation of cosmid and cDNA clones in the region surrounding the
RT BTK gene at Xq21.3-q22.";
RL Genomics 21:517-524(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Oelton J.C., Liu X., Lu J., Malley T.M., Allen R.C., Muzny D.M.,
RA Belmont J.W., Gibbs R.A.;
RN Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
RN [3]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=96081943; PubMed=7499401;
RA Honore B., Rasmussen H.H., Vorum H., Dejgaard K., Liu X.,
RA Gromov P., Madsen P., Gesser B., Tommerup N., Celis J.E.;
RT "Heterogeneous nuclear ribonucleoproteins H, H', and F are members of
RT a ubiquitously expressed subfamily of related but distinct proteins
RT encoded by genes mapping to different chromosomes.";
RL J. Biol. Chem. 270:28780-28789(1995).
RN [4]
RP FUNCTION: THIS PROTEIN IS A COMPONENT OF THE HETEROGENEOUS NUCLEAR
RN RIBONUCLEOPROTEIN (HNRNP) COMPLEXES WHICH PROVIDE THE SUBSTRATE
RN FOR THE PROCESSING EVENTS THAT PRE-MRNAS UNDERGO BEFORE BECOMING
RN FUNCTIONAL, TRANSLATABLE MRNAS IN THE CYTOPLASM. BINDS POLY (RG).
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOPLASM.
CC -1- TISSUE SPECIFICITY: EXPRESSED UBQUITOUSLY.
CC -1- SIMILARITY: CONTAINS 3 RNA RECOGNITION MOTIFS (RRM).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U01923; -; NOT_ANNOTATED_CDS.

DR EMBL: U78027; AAB64202.1; -.
DR Aarhus/Chent-2DPAGE; 4432; IEF.
DR MIM: 601036; -.
DR InterPro; IPR000504; -.
DR Pfam; PF00076; rrm; 3.
DR PROSITE; PS50102; RRM; 3.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat.
FT DOMAIN 11 90 RNA-BINDING (RRM) 1.
FT DOMAIN 111 188 RNA-BINDING (RRM) 2.
FT DOMAIN 289 364 RNA-BINDING (RRM) 3.
FT DOMAIN 234 433 2 X 16 AA GLY-RICH APPROXIMATE REPEATS.
FT REPEAT 234 249 1-1.
FT REPEAT 418 433 1-2.
FT DOMAIN 354 392 2 X 19 AA PERFECT REPEATS.
FT REPEAT 354 372 2-1.
FT REPEAT 374 392 2-2.
SQ SEQUENCE 449 AA; 49263 MW; C892523A638F07C7 CRC64;

Query Match 69.6%; Score 32; DB 1; Length 449;
Best local Similarity 85.7%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RIGHTI 8
| | | | |
DB 175 RIGHTI 181

Search completed: June 20, 2001, 14:09:31
Job time: 328 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2001, 14:04:38 ; Search time 47.6 Seconds

(without alignments)
25.016 Million cell updates/sec

Title: US-09-692-401-4

Perfect score: 46

Sequence: 1 VRIGHLYIL 9

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_PROTOCTA:*
12: SP_UNCLASSIFIED:*
13: SP_VERTEBRATE:*
14: SP_VIRUS:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	314	4 Q9NSD3	Q9nsd3 homo sapien
2	35	76.1	233	10 Q9S181	Q9s181 arabidopsis
3	35	76.1	234	2 Q9S1R5	Q9s1r5 streptomyce
4	35	76.1	305	2 Q45454	Q45454 bacillus su
5	35	76.1	520	10 Q65815	Q65815 helianthus
6	35	76.1	531	2 Q9RM20	Q9rm20 bacillus an
7	35	76.1	1310	5 Q9VVP8	Q9vvp8 drosophila
8	34	73.9	139	2 Q49122	Q49122 methylobact
9	34	73.9	637	2 Q9EVH2	Q9evh2 buchnera ap
10	34	73.9	637	2 Q06769	Q06769 mycobacteri
11	34	73.9	723	4 Q9UD06	Q9ud06 homo sapien
12	34	73.9	1244	10 Q82607	Q82607 arabidopsis
13	34	73.9	1461	10 Q9S1M3	Q9s1m3 arabidopsis
14	33	71.7	313	1 Q28287	Q28287 archaetoglob
15	33	71.7	324	2 Q924K0	Q924k0 streptomyce
16	33	71.7	337	1 Q9YFK2	Q9yfk2 aeropyrum p
17	33	71.7	338	1 Q9UWX3	Q9uwx3 sulfobolus
18	33	71.7	423	10 Q49639	Q49639 arabidopsis
19	33	71.7	444	2 Q85748	Q85748 treponema d

20	32	69.6	154	6 Q9M2A9	Q9m2a9 ovis aries
21	32	69.6	232	14 Q9QP22	Q9qp22 gallid heyr
22	32	69.6	297	4 Q9NP96	Q9np96 homo sapien
23	32	69.6	301	2 Q52770	Q52770 pseudomonas
24	32	69.6	317	2 Q9L2G1	Q9l2g1 streptomyce
25	32	69.6	325	2 Q34558	Q34558 bacillus su
26	32	69.6	325	9 Q64101	Q64101 bacterioph
27	32	69.6	325	10 Q43323	Q43323 saccharum h
28	32	69.6	331	4 Q9NP47	Q9np47 homo sapien
29	32	69.6	346	4 Q9Y4J5	Q9y4j5 homo sapien
30	32	69.6	350	2 Q9W292	Q9w292 thermotoga
31	32	69.6	354	10 Q9SYZ8	Q9syz8 arabidopsis
32	32	69.6	415	11 Q9Z2X1	Q9z2x1 rattus norv
33	32	69.6	442	4 Q9NWX1	Q9nwx1 homo sapien
34	32	69.6	449	5 Q23284	Q23284 caenorhadi
35	32	69.6	449	11 Q35737	Q35737 mus musculu
36	32	69.6	449	11 P70333	P70333 mus musculu
37	32	69.6	657	4 Q9UJ28	Q9uj28 homo sapien
38	32	69.6	657	11 Q9R0C5	Q9r0c5 rattus norv
39	32	69.6	666	5 Q9VZX5	Q9vzx5 drosophila
40	32	69.6	889	10 Q9FXA5	Q9fxa5 arabidopsis
41	32	69.6	1845	14 Q9WJZ8	Q9wjz8 human cytom
42	32	69.6	1863	11 Q9J0L1	Q9j0l1 mus musculu
43	31.5	68.5	336	1 Q9HKS5	Q9hks5 thermoplas
44	31	67.4	101	2 Q9PH51	Q9ph51 xyella tas
45	31	67.4	146	14 Q9J511	Q9j511 fowlpox vir

ALIGNMENTS

RESULT 1	
Q9NSD3	PRELIMINARY; PRT; 314 AA.
ID Q9NSD3;	
AC Q9NSD3;	
DT 01-OCT-2000 (TREMBLrel. 15, Created)	
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)	
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
DE MELANOMA ANTIGEN FAMILY A12.	
GN MAGEA12.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
OX NCBI_TaxID=9606;	
RN [1]	
RP SEQUENCE FROM N.A.	
RA Mellon A.M., Platzer M., Bates R., Gloeckner G., Botcherby M.,	
RA Nordstiek G., Striyens M.A., Kioschis P., Dangel A., Cunningham D.,	
RA Straw R., Weston P., Hunter C., Gilbert M., Fernando S., Goodall K.,	
RA Kerry G., Greystrom J.S., Clark D., Goerdes M., Blechschmidt K.,	
RA Rump A., Hinzmann B., Mundy C.R., Miller W., Poustka A., Herman G.E.,	
RA Rhodes M., Denny P., Rosenthal A., Brown S.D.M.;	
RT "Comparative genome sequence analysis of the Bpa/Str region in mouse	
RT and man.;"	
RL Genome Res. 0:0-0(2000).	
DR EMBL: U82671; AAF44789.1; .	
DR InterPro: IPR002190; .	
DR Pfam: PF01454; MAGE; 1.	
SQ SEQUENCE 314 AA; 34836 MW; 7E00F7CECD8F6568 CRC64;	
Query Match	100.0%; Score 46; DB 4; Length 314;
Best local Similarity	100.0%; Pred. No. 0.32;
Matches	9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VRIGHLYIL 9	
DB 170 VRIGHLYIL 178	
RESULT 2	
Q9S181	PRELIMINARY; PRT; 233 AA.
ID Q9S181	

AC 09SIR1:
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE F23N19.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Walker M., Shih P., Brooks S., Buehler E., Chao Q., Dunn P., Khan S.,
RA Kimer M., Altati H., Araujo R., Conn L., Conway A.B., Gonzalez A.,
RA Hansen N.F., Hultzer L., Kremetska I., Lenz G., Li J., Liu S.,
RA Lucos S., Rowley D., Schwartz J., Toriumi M., Vysotskaia V., Yu G.,
RA Davis R.W., Federspiel N.A., Theologis A., Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F23N19 from chromosome
RT 1".
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC007190; AAF19536.1; -.
SQ SEQUENCE 233 AA; 26150 MW; 7D292F68D261C6EE CRC64;

Query Match 76.1%; Score 35; DB 10; Length 233;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VRIGHLY 7
Db 177 VRIGHVY 183

RESULT 3
Q9SIR5 PRELIMINARY; PRT; 234 AA.
ID Q9SIR5;
AC Q9SIR5;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE PUTATIVE MEMBRANE PROTEIN.
GN SCJ9A.05C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetales; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-A3(2);
RA Harris D.;
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN-A3(2);
RC MEDLINE=97000351; PubMed=8843436;
RX Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL: AL109972; CAB53266.1; -.
SQ SEQUENCE 234 AA; 23619 MW; AAFEB11CD014319 CRC64;

Query Match 76.1%; Score 35; DB 2; Length 234;
Best Local Similarity 55.6%; Pred. No. 36;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VRIGHLYL 9
Db 148 LEIGHMYL 156

RESULT 4
Q45454 PRELIMINARY; PRT; 305 AA.
ID Q45454;
AC Q45454;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE PROBABLY POSITIONED IN OPERON WITH DOWNSTREAM ORFs.
OS Bacillus subtilis.
OC Bacteri; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO3022;
RX MEDLINE=95206941; PubMed=7899081;
RA Meljer W.J., Venema G., Bron S.;
RT "Characterization of single strand origins of cryptic rolling-circle
RT plasmids from Bacillus subtilis.";
RL Nucleic Acids Res. 23:612-619(1995).
DR EMBL: U32380; AAC44422.1; -.
RW Plasmid.
SQ SEQUENCE 305 AA; 35592 MW; DBA10B2A7BC51B22 CRC64;

Query Match 76.1%; Score 35; DB 2; Length 305;
Best Local Similarity 66.7%; Pred. No. 47;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 VRIGHLYL 9
Db 197 VLIHMYVL 205

RESULT 5
O65815 PRELIMINARY; PRT; 520 AA.
ID O65815;
AC O65815;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE CYTOCHROME P450 81B1 (EC 1.14.13.53) (ISOFLAVONE 2'-HYDROXYLASE).
GN CYP81B1 OR CYP81B1L OR CYP81B1C OR CYP81B1S.
OS Helianthus tuberosus (Jerusalem artichoke).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asterales; Heliantheae;
OC eunasterids II; Asterales; Asteraceae; Asteroideae; Helianthus.
OX NCBI_TaxID=4233;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. BLANC COMMUN; TISSUE=TUBER;
RX MEDLINE=98184826; PubMed=9516419;
RX Capello-Hurtado F., Batard Y., Salau J.-P., Durst F., Pinot F.,
RA Weick-Reichardt D.;
RT "Cloning, expression in yeast, and functional characterization of
RT CYP81B1, a plant cytochrome P450 that catalyzes in-chain hydroxylation
RT of fatty acids";
RL J. Biol. Chem. 273:7260-7267(1998).
CC -1- FUNCTION: CATALYZES THE HYDROXYLATION OF ISOFLAVONES, DAIDZEIN AND
CC FORMONONETIN, TO YIELD 2'-HYDROXYISOFLAVONES, 2'-HYDROXYDAIDZEIN,
CC AND 2'-HYDROXYFORMONONETIN, RESPECTIVELY (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: FORMONONETIN + NADPH + O(2) =
CC 2'-HYDROXYFORMONONETIN + NADP(+) + H(2)O.
CC -1- COFACTOR: THE ENZYME IS NADPH-DEPENDENT.
CC -1- ENZYME REGULATION: ENZYME ACTIVITY IS INDUCED BY MNCL2 OR
CC AMINOPURINE.

CC -1- PATHWAY: INVOLVED IN THE BIOSYNTHESIS OF ISOFLAVONOID-DERIVED
 CC ANTIMICROBIAL COMPOUNDS OF LEGUMES.
 CC -1- ALTERNATIVE PRODUCTS: TWO FORMS: CYP81B1L (SHOWN HERE) AND
 CC CYP81B1S: ARE PRODUCED BY ALTERNATIVE SPLICING
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL: AJ000478; CAA04117.1; -
 DR EMBL: AJ000477; CAA04116.1; -
 DR InterPro: IPR001128; -
 DR Pfam: PF00067; P450.1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
 KM Oxidoreductase; Monooxygenase; Membrane; Heme; Alternative splicing;
 KM NAD.
 FT BINDING 456 456 HEME (BY SIMILARITY).
 FT VARSPLC 9 22 MISSING (IN CYP81B1S).
 FT CONFLICT 40 40 MISSING (IN CAA04116).
 FT CONFLICT 51 51 Y->S (IN CAA04116).
 FT CONFLICT 78 78 P->O (IN CAA04116).
 SQ SEQUENCE 520 AA; 58914 MW; 38108E379BDB35C8 CRC64;

Query Match 76.1%; Score 35; DB 10; Length 520;
 Best Local Similarity 85.7%; Pred. No. 83;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 IGHLYL 9
 |||||:
 DB 55 IGHLYL 61

RESULT 6
 GORMZ0
 ID O9RMZ0 PRELIMINARY; PRT; 531 AA.
 DT 01-MAY-2000 (TREMBLrel. 13. Created)
 DT 01-MAY-2000 (TREMBLrel. 13. Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16. Last annotation update)
 DE PXO2-42.
 OS Bacillus anthracis.
 OC Plasmid PXO2.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1392;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Okinaka R.T., Cloud K., Hampton O., Hill K.K., Keim P., Lanke G.,
 RA Kuno S.E., Manter D., Martinez Y., Svensson R., Tatum L.R.,
 RA Brown A.E., Jackson P.J.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF188935; AAF13647.1; -
 DR InterPro: IPR001119; -
 DR InterPro: IPR002508; -
 DR Pfam: PF00395; SLH; 3.
 DR Pfam: PF01520; Amidase_3; 1.
 KM Plasmid
 SQ SEQUENCE 531 AA; 58864 MW; 9A171DC4ED05CA78 CRC64;

Query Match 76.1%; Score 35; DB 2; Length 531;
 Best Local Similarity 66.7%; Pred. No. 85;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRIIGHLYL 9
 |::|::|:
 DB 467 VRIIGHLYL 475

RESULT 7
 G9VVF8 PRELIMINARY; PRT; 1310 AA.
 ID G9VVF8
 AC G9VVF8
 DT 01-MAY-2000 (TREMBLrel. 13. Created)
 DT 01-MAY-2000 (TREMBLrel. 13. Last sequence update)

DT 01-MAY-2000 (TREMBLrel. 13. Last annotation update)
 DE CG7692 PROTEIN.
 GN CG7692.
 OS Drosophila melanogaster (Fruit fly)
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brostein P., Brotier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.D., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ideyama C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mantei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Nobarray C., Morris J., Moshell A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington S., Saunders R.D.C., Scheefer F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003524; AAF49353.1; -
 DR FlyBase: FBgn0036714; CG7692.
 SQ SEQUENCE 1310 AA; 150503 MW; 2AF07754E8452430 CRC64;

Query Match 76.1%; Score 35; DB 5; Length 1310;
 Best Local Similarity 85.7%; Pred. No. 2.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 IGHLYL 9
 |||||:
 DB 219 IGHLYL 225

RESULT 8
 Q49122 PRELIMINARY; PRT; 139 AA.
 ID Q49122
 AC Q49122;
 DT 01-NOV-1996 (TREMBLrel. 01. Created)
 DT 01-NOV-1996 (TREMBLrel. 01. Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16. Last annotation update)
 DE MAU GENE CLUSTER. METHYLAMINE DEHYDROGENASE LARGE AND SMALL SUBUNITS,
 DE AND AMICYANIN. (MAUPBEDACJGLMN) GENES, COMPLETE CDS (FRAGMENT).

```

OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=408;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AM1:
RA MEDLINE=94292425; PubMed=8021187;
RA Christoserdov A.Y., Christoserdova L.V., McIntire W.S., Lidstrom M.E.;
RT "Genetic organization of the mau gene cluster in Methylobacterium
RT extorquens AM1: complete nucleotide sequence and generation and
RT characteristics of mau mutants.";
RL J. Bacteriol. 176:4052-4065(1994).
DR EMBL: L26406; AAB46931.1; -
DR InterPro: IPR001387; -
DR Pfam: PF01381; HTH_3; 1.
DR SMART: SM00530; HTH_XRE; 1.
FT NON_TER 139
SQ SEQUENCE 139 AA; 14906 MW; 5CEBA4626D3183B1 CRC64;

Query Match 73.9%; Score 34; DB 2; Length 139;
Best Local Similarity 55.6%; Pred. No. 33;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VRIGHLYL 9
Db 104 IKVGHLYL 112

RESULT 9
O9EVH2 PRELIMINARY; PRT; 363 AA.
AC O9EVH2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE 3-ISOPROPYLMALATE DEHYDROGENASE.
GN LEBU.
OS Buchnera aphidicola.
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=9;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-UCAL:
RA Wernegreen J.J., Moran N.A.;
RT "Vertical Transmission of Biosynthetic Plasmids in Aphid Symbionts
RT (Buchnera)".;
RL J. Bacteriol. 183:0-0(2001).
DR EMBL: AF197453; AAG31393.1; -
KW Plasmid.
SQ SEQUENCE 363 AA; 40446 MW; 44A79DE5ADFCE243 CRC64;

Query Match 73.9%; Score 34; DB 2; Length 363;
Best Local Similarity 75.0%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VRIGHLYL 8
Db 219 VRLSHLYL 226

RESULT 10
O06769 PRELIMINARY; PRT; 637 AA.
AC O06769;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE HYPOTHETICAL 69.5 KDA PROTEIN.
GN RV0669C OR MTI376.05.

```

```

OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RA MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Chillingworth T., Connor R.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
RA Honnsby T., Jagsels K., Kirogh A., Mclean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
DR EMBL: Z95972; CAB09388.1; -
DR Tuberculist; RV0669C; -
KW Hypothetical protein.
SQ SEQUENCE 637 AA; 69489 MW; 5E2D915AB1E4FDB7 CRC64;

Query Match 73.9%; Score 34; DB 2; Length 637;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VRIGHLYL 9
Db 404 VRIGHLYL 412

RESULT 11
Q9UDU6 PRELIMINARY; PRT; 723 AA.
AC Q9UDU6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE HEPATOCYTE GROWTH FACTOR.
GN HGF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Courtney L., Elliot G., Angell S.;
RT "The Sequence of Homo sapiens PAC clone RPS-1098B1.".
RL Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Waterston R.H.;
RL Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC -1- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR EMBL: AC004960; AAC71655.1; -
DR HSSP; P14210; 1BHT.
DR InterPro: IPR000001; -
DR InterPro: IPR000327; -
DR InterPro: IPR001254; -
DR InterPro: IPR001314; -
DR InterPro: IPR003014; -
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00051; kringleg; 4.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.

```

PRINTS; PRO0018; KRINGLE.
 DR PRODM; PD000583; -, 1.
 DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_4.
 DR PROSITE; PS50070; KRINGLE_2; 4.
 DR SMART; SM00020; TRYP_Spc1.1.
 KM Hydrolase; Serine protease.
 SO SEQUENCE 723 AA; 82602 MW; 627B1EF99FAD931B CRC64;

Query Match 73.9%; Score 34; DB 4; Length 723;
 Best Local Similarity 55.6%; Pred. No. 1.9e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VRIHXYTL 9
 :|:|:|:
 Db 624 LRVAHLYIM 632

RESULT 12
 ID 082607 PRELIMINARY; PRT; 1244 AA.
 AC 082607;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
 DE T2L5.9 PROTEIN.
 GN T2L5.9.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurossids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA WASHU;
 RT "The A. thaliana Genome Sequencing Project."
 RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Geisel C., Smith A., Le T.;
 RT "The sequence of A. thaliana T2L5."
 RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Waterston R.;
 RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
 CC -1 SIMILARITY: BELONGS TO ZN-FINGER CCHC TYPE FAMILY.
 DR EMBL; AF096371; AAC62795.1; -.
 DR InterPro; IPR001584; -.
 DR InterPro; IPR001878; -.
 DR Pfam; PF00098; zf-CCHC; 1.
 DR Pfam; PF00665; rve; 1.
 DR PRINTS; PR00939; C2HC2NFNGSR.
 DR SMART; SM00343; znf_C2HC; 1.
 KM Zinc-finger.
 SO SEQUENCE 1244 AA; 13937 MW; C2B34AA509B5C80 CRC64;

QY 2 RIGHLYTL 9
 |||:|:|:
 Db 520 RIGMLYVL 527

Query Match 73.9%; Score 34; DB 10; Length 1244;
 Best Local Similarity 75.0%; Pred. No. 3.3e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

RESULT 13
 Q9S1M3 PRELIMINARY; PRT; 1461 AA.
 ID Q9S1M3

AC Q9S1M3;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
 DE PUTATIVE RETROELEMENT POLYPROTEIN.
 GN ATG20460.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurossids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldguy T.V.,
 Buel C.R., Ketchum K.A., Lee J.J., Rongling C.M., Koo H., Moffat K.S.,
 Cronin L.A., Shen M., VanAken S.E., Umeyam L., Tallon L.J., Gill J.E.,
 Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 R. Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome II of Arabidopsis thaliana."
 CC Nature 402:761-768(1999).
 DR EMBL; AC007109; AAD23646.1; -.
 DR InterPro; IPR001584; -.
 DR InterPro; IPR001878; -.
 DR Pfam; PF00098; zf-CCHC; 1.
 DR Pfam; PF00665; rve; 1.
 DR SMART; SM00343; znf_C2HC; 1.
 KM Polyprotein; Zinc-finger.
 SO SEQUENCE 1461 AA; 162670 MW; E79E2C46BBE3D16 CRC64;

Query Match 73.9%; Score 34; DB 10; Length 1461;
 Best Local Similarity 75.0%; Pred. No. 3.9e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 RIGHLYTL 9
 |||:|:|:
 Db 544 RIGMLYVL 551

RESULT 14
 ID 028287 PRELIMINARY; PRT; 313 AA.
 AC 028287;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
 DE CALCIUM-BINDING PROTEIN, PUTATIVE.
 GN AFI92.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 OC Archaeoglobus.
 OC NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
 Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 Kitzness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 Peterson S., Reich C.T., McNeil L.K., Badger J.H., Glöck A., Zhou L.,
 Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Usterback T.,
 Cotton M.D., Spriggs T., Arlisch P., Kaine B.P., Sykes S.M.,
 Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-

RF reducing archaean Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 DR EMBL; AE000965; AAB89260.1; .
 DR TIGR; AF1992; .
 DR InterPro; IPR002048; .
 DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 313 AA; 34574 MW; 4FDD50CB5BBA4A CRC64;

Query Match 71.7%; Score 33; DB 1; Length 313;
 Best Local Similarity 62.5%; Pred. No. 1.2e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VRIGLXYI 8
 :|||:|
 DB 272 IEIGHYVI 279

RESULT 15
 O924K0 PRELIMINARY; PRT; 324 AA.
 AC O924K0:
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)
 DE FACTOR C PROTEIN PRECURSOR.
 OS Streptomyces griseus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1911;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=45H;
 RX MEDLINE=80156796; PubMed=6767606;
 RA Biro S., Bekesi I., Vitalis S., Szabo G.;
 RT "A substance effecting differentiation in Streptomyces griseus.
 RT Purification and properties.";
 RL Eur. J. Biochem. 103:359-363(1980).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=45H;
 RA Biro Z., Smegei A., Szeszak F., Vitalis S., Szabo G., Biro S.;
 RT "Cloning and sequencing of the factor C gene from Streptomyces griseus
 RT 45H.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF103943; AAC97368.1; .
 KW Signal.
 FT SIGNAL. 1 38 POTENTIAL.
 FT CHAIN 39 324 FACTOR C PROTEIN.
 SQ SEQUENCE 324 AA; 34555 MW; E5F57BA09D2BAEDB CRC64;

Query Match 71.7%; Score 33; DB 2; Length 324;
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 RIGHLYIL 9
 :|||:|
 DB 106 KIGHMYLL 113

Search completed: June 20, 2001, 14:10:22
 Job time: 344 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2001, 14:03:13 ; Search time 41.58 Seconds

(Without alignments)
12.373 Million cell updates/sec

Title: US-09-692-401-4
Perfect score: 46
Sequence: 1 VRIGHLYIL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	82.6	9	15	R50283
2	38	82.6	9	17	R99345
3	38	82.6	9	20	Y46062
4	38	82.6	10	18	W23043
5	35	76.1	505	20	Y05802
6	35	76.1	621	20	Y34971
7	34	73.9	189	13	R25692
8	34	73.9	697	21	Y98485
9	34	73.9	697	21	Y59030
10	34	73.9	697	22	B45838
11	34	73.9	723	11	R07144

12	34	73.9	723	12	R14307	Plasminogen-like g
13	34	73.9	723	12	R15624	Human leukocyte-de
14	34	73.9	723	13	R21142	Human TCF-II. Hom
15	34	73.9	723	13	R25677	Recombinant human
16	34	73.9	723	13	R29819	TCF-II. R29819;
17	34	73.9	723	15	R57026	Human wild-type tu
18	34	73.9	723	15	R57027	Human modified tum
19	34	73.9	723	15	R57028	Human modified tum
20	34	73.9	723	16	R82685	Tumour cytotoxic f
21	34	73.9	723	17	R89688	TCF mutant having
22	34	73.9	723	17	R99689	Human plasminogen-
23	34	73.9	723	19	W76690	Human leukocyte-de
24	34	73.9	723	19	W59923	Hepatic parenchyma
25	34	73.9	727	12	R10656	Human hepatocyte g
26	34	73.9	728	12	R12792	Human hepatocyte g
27	34	73.9	728	12	R14243	Human hepatocyte g
28	34	73.9	728	12	R15623	Human hepatocyte g
29	34	73.9	728	13	R20005	Human hepatocyte g
30	34	73.9	728	13	R20100	Contains alpha - an
31	34	73.9	728	13	R21976	Human Hepatocyte g
32	34	73.9	728	13	R25160	Human HGF. Homo s
33	34	73.9	728	13	R25676	Recombinant humo s
34	34	73.9	728	14	R40862	Competitive inhibi
35	34	73.9	728	14	R40863	Competitive inhibi
36	34	73.9	728	14	R39521	Hepatocyte growth
37	34	73.9	728	14	R42062	Vascular endotheli
38	34	73.9	728	14	R52940	Human hepatocyte g
39	34	73.9	728	14	R52941	Human hepatocyte g
40	34	73.9	728	14	R52942	Human hepatocyte g
41	34	73.9	728	14	R52943	Human hepatocyte g
42	34	73.9	728	14	R52944	Human hepatocyte g
43	34	73.9	728	14	R52945	Human hepatocyte g
44	34	73.9	728	14	R52946	Human hepatocyte g
45	34	73.9	728	14	R52947	Human hepatocyte g

ALIGNMENTS

RESULT	1	
ID	R50283	standard; Protein: 9 AA.
XX	AC	R50283;
XX	AC	26-SEP-1994 (first entry)
DT	XX	MAGE-21 nonapeptide.
XX	XX	
DE	XX	MAGE, nonapeptide; cancer; melanoma; breast cancer; HLA.
KW	KW	histocompatibility; human leucocyte antigen; probe; treatment;
KW	KW	therapy; vaccine.
XX	XX	
OS	OS	Synthetic.
XX	XX	
PN	PN	W09405304-A.
XX	XX	
PD	PD	17-MAR-1994.
XX	XX	
PF	PF	30-AUG-1993; 93WO-US08157.
XX	XX	
PR	PR	31-AUG-1992; 92US-0938334.
PR	PR	26-MAR-1993; 93US-0037230.
XX	XX	07-JUN-1993; 93US-0073103.
XX	XX	
FA	FA	(LUDW-) LUDWIG INST CANCER RES.
XX	XX	
PI	PI	Boon-falleur T, De Plaen E, Lurquin C, Traversari C;
PI	PI	Van Derbruggen P;
XX	XX	
DR	DR	WPI: 1994-100844/12.
XX	XX	N-PSDB; Q44753.

PT New nona:peptide derived from tumour rejection antigen precursor
PT - presented by HLA-A1 cancer cells, for use in diagnosis or
PT therapy of esp. melanoma and breast cancer.
PS
PS Disclosure; Page 20; 33pp; English.
XX
XX An isolated nonapeptide having the amino acid sequence Glu-Val-Asp-
CC Pro-Ile-Gly-His-Leu-Tyr is derived from the tumour rejection antigen
CC precursor encoded by the MAGE-3 gene and presented by HLA-A1. The
CC nonapeptide can be used in a vaccine to treat a cancerous condition
CC involving HLA-A1 subtype cancerous cells. The nucleic acid encoding
CC the nonapeptide can be used as a probe to identify tumour cells.
CC This sequence is homologous to the peptide described and is encoded
CC by the MAGE-21 gene.
XX
XX
SQ Sequence 9 AA;

Query Match 82.6%; Score 38; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRIGHLY 7
| | | | |
Db 3 vrighly 9

RESULT 2

R99345 ID R99345 standard; Protein; 9 AA.

AC R99345;

DT 22-APR-1997 (first entry)

DE MAGE-21 nonapeptide.

XX HLA binding peptide; cell lysis; cytolytic T cell; MAGE family; human;
KW tumour rejection antigen precursor; TRA; MAGE-1; tumour; cancer cell;
KW antibody; melanoma; universal effector cell; vaccine; breast cancer; CTL;
KW therapy.
XX
XX Homo sapiens.

PN W09626214-A1.

PD 29-AUG-1996.

PF 01-FEB-1996; 96WO-US01489.

PR 23-FEB-1995; 95US-0393273.

PA (LUDWIG) LUDWIG INST CANCER RES.

PI Boon-Falleur T, De Plaen E, Gaugler B, Lurquin C;

PI Romero P, Traversari C, Van Den Eynde B, Van Der Bruggen P;

DR WPI; 1996-402317/40.

DR N-PSDB; T35410.

XX New nona:peptide(s) that bind to HLA molecule(s) and induce lysis
PT by specific cytolytic T cells, for diagnosis and treatment of
PT tumours and to expand T cells in vitro.
XX
XX

PS Example 4; Fig 4; 41pp; English.

CC R99343-R99350 represent MAGE nonapeptides, based on the tumour rejection
CC antigen region of the full length MAGE sequences. These peptides were
CC used to design the nonapeptides of the invention (see R99337-R99342),
CC which bind to a HLA molecule on a cell, and provoke lysis by cytolytic T
CC cells (CTLs) specific for a complex of the HLA molecule and nonapeptide.
CC The nonapeptides can be used diagnostically to identify tumours
CC expressing a particular HLA molecule, or to identify cancer cells. The

CC peptides can also be used therapeutically, to induce a CTL response to
CC tumours (where the peptides are optionally coupled to tumour-specific
CC antibodies), or to induce a response by CTLs that are otherwise inactive.
CC The peptide sequences may also be used to expand specific CTLs in vitro
CC for later return to the patient, such as for treating melanoma. Tumour
CC cells can be identified by using DNA encoding the nonapeptides as probes.
CC Non-human cells transformed with the HLA-A1 gene and a DNA sequence
CC encoding one of the peptides, can be used to generate CTLs, or to detect
CC the presence of CTLs in human samples. The non-human transformed cells,
CC when polytransformed, are universal effector cells, and can be used in
CC vaccines, or for treating melanoma or breast cancer.
XX
XX
SQ Sequence 9 AA;

Query Match 82.6%; Score 38; DB 17; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRIGHLY 7
| | | | |
Db 3 vrighly 9

RESULT 3

Y46062 ID Y46062 standard; Peptide; 9 AA.

AC Y46062;

DT 01-DEC-1999 (first entry)

DE Immunogenic peptide having a human leukocyte antigen binding motif #673.

XX Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW immune response; T cell activation; major histocompatibility complex;
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW vaccine; immunisation.
XX
XX Synthetic.

OS Homo sapiens.

PN W09945954-A1.

PD 16-SEP-1999.

PF 13-MAR-1998; 98WO-US05039.

PR 13-MAR-1998; 98WO-US05039.

PA (EPIM-) EPIMUNE INC.

PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;

DR WPI; 1999-551214/46.

XX New immunogenic peptides with HLA binding motif, useful in treatment
PT and diagnosis of cancers and viral diseases
PT
XX

PS Claim 1; Page 54; 150pp; English.

CC Y45390 to Y48214 represent specifically claimed immunogenic peptides
CC having a human major histocompatibility complex (MHC) Class I (also
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC response against the antigen from which the peptide is derived.
CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC normally induced by an antigen in the form of a peptide fragment bound
CC to a HLA molecule, rather than the intact foreign antigen itself, and
CC are particularly important in tumour rejection and in fighting viral
CC infections. The peptides are therefore useful therapeutically to treat

CC or prevent viral infections and cancers in mammals (especially humans
CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
CC They can be administered as vaccines to elicit an immune response in
CC individuals susceptible or otherwise at risk of viral infection or
CC cancer, or used to treat chronic or acute conditions. They are also
CC useful diagnostically, and can be used to induce a cytotoxic T cell
CC response, by contacting a cytotoxic T cell with the peptide e.g. to
CC produce CTLs ex vivo for infusion back into a patient. The
CC polypeptides encoding the immunogenic peptides are also useful
CC therapeutically and for immunisation as above.
CC
CC
CC
CC Sequence 9 AA:
CC
CC
CC

Query Match	82.6%;	Score 38;	DB 20;	Length 9;
Best Local Similarity	100.0%;	Pred. No. 3.2e+05;		
Matches	7;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0

QY	1	VRIGLY	7
Db	3	vrighly	9

RESULT	4
W23043	
ID	W23043 standard; Peptide; 10 AA

AC W23043
XX
DT 25-FEB-

DT	25-FEB-1998	(first entry)
XX		
DE	MAGE-12/HLA-B44	tumour rejection antigen

KW MAGE-12; tumour rejection antigen precursor; TRAP; HLA-B44;
KW human leukocyte antigen B44; cytotoxic T lymphocyte; cancer;
KW melanoma; therapy; diagnosis; vaccine.

OS	Homo sapiens.
XX	
PN	W09731017-A1.

PD	28-AUG-1997.
XX	
PF	05-FEB-1997;

PR 20-FEB-1996; 96US-0602506.
XX
PA (LUDW-) LUDWIG INST CANCER RES

PI Boon-Falleur T, Coulle P, Herman J, Luescher I
PI Van Der Bruggen P;
XX

XX Tumour rejection antigens presented by human leukocyte antigen B44
PT molecules - useful to identify HLA-B44 positive cells for diagnosis
PT

XX
PS
XX

Claim 2; Page 51; 74pp; English

CC This peptide is a tumour rejection antigen presented by a HLA-B*44
CC molecule and derived from a MAGE-12 tumour rejection antigen
CC precursor (TRAP). Claimed tumour rejection antigens (#20308-43)
CC are able to bind to HLA-B*44 positive cells, making them useful in
CC identifying cells which present HLA-B*44 molecules on their
CC surfaces for use in the diagnosis and therapy of cellular
CC abnormalities. The complex of the tumour rejection antigen and HLA
CC molecule provides a cytolytic T cell response. The tumour
CC rejection antigens, or complexes of tumour rejection antigens and
CC HLA-B*44, can be used as vaccines to treat disorders characterised
CC by expression of the TRAP molecule such as cancer, especially
CC melanoma. Vaccines can also be prepared from cells which present
CC the tumour rejection antigen/HLA complexes on their surface, such

CC as non-proliferative cancer cells and non-proliferative
CC transfectants.
XX
SQ Sequence 10 AA;

Sequence	10 AA;
SQ	

Query Match	82.6%;	Score 38;	DB 18;	Length 10;					
Best Local Similarity	100.0%;	Pred. No. 0.099;							
Matches	7;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0

QY	1	VRIGLY	7
Db	4	vriightly	10

RESULT	5
Y05902	
ID	Y05902 standard; Protein; 505 AA

AC	Y05902;
XX	
DT	02-AUG-1999 (first entry)

DE Jerusalem artichoke in-chain hydroxylase CYP81B1
XX
KW In-chain hydroxylase; transgenic plant; lipid; hy

OS	<i>Helianthus tuberosus</i> .
XX	
KW	CYP81B1; cytochrome P450.
AS	

PN	WO9918224-A1.
XX	
PD	15-APR-1999.

PF	06-OCT-1998;	98WO-IB01716.
XX		
PR	06-OCT-1997;	97US-0060960.

PA (CNRS) CENT NAT RECH SCI.
XX
PI Batard Y, Benveniste I, Cabello-Huartado F, Durst F,

PI Werck-Reichhart D;
XX
DR WPI; 1999-264030/22.

XX Nucleic acid encoding plant fatty acid hydroxylases

XX The present sequence represents in-ch
CC Jerusalem artichoke. CYP81B1 is a mi
CC

CC catalyses the omega-2, omega-3 and omega-4 hydroxylation of capric,
CC lauric and myristic acids. The major metabolite is the
CC omega-3-hydroxylated compound. The invention provides isolated
CC nucleic acids (see X58400-06), encoding plant fatty acid
CC hydroxylases (see Y05895-902). Also claimed are host cells,
CC transgenic plants and compositions consisting of the plant fatty
CC acid hydroxylase, a process for isolating additional fatty acid
CC hydroxylase genes from a plant, and a process of altering the fatty
CC acid composition in a plant by expressing the plant fatty acid
CC hydroxylase in a transgenic plant, and hydroxylating or epoxidizing
CC a fatty acid substrate in the plant. Manipulating the hydroxylated
CC fatty acid content of plants will modify resistance to drought and
CC attack by insects and other pests. The transgenic plants may also
CC be used as sources of hydroxylated and epoxidized fatty acids
CC useful in the manufacture of e.g. lubricants, anti-slip agents,
CC plasticisers, coating agents, detergents and surfactants.

Query Match 76.1%; Score 35; DB 20; Length 505;
 Best Local Similarity 85.7%; Pred. No. 31;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 IGHLYL 9
 |||||:
 DB 40 Ighlyl 46

RESULT 6

Y34971
 ID Y34971 standard; Protein; 621 AA.

XX Y34971;

DT 13-SEP-1999 (first entry)

DE Amino acid sequence of a Chlamydia pneumoniae protein.

KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
 KW vaccine; neutralising epitope.

OS Chlamydia pneumoniae.

PN MO9927105-A2.

PD 03-JUN-1999.

PF 20-NOV-1998; 98WO-IB01890.

PR 04-NOV-1998; 98US-0107078.

PR 21-NOV-1997; 97FR-0014673.

XX (GSEST) GENSET.

PI Griffiths R;

DR WPI; 1999-357842/30.

PT Genome sequence of Chlamydia pneumoniae

PS Page 885-886; Disclosure; 1912pp; English.

CC Y34584-Y35879 represent the proteins encoded by all the open reading
 CC frames in the complete genome (see X91990) of Chlamydia pneumoniae.
 CC C. pneumoniae causes respiratory disease such as pneumonia and
 CC bronchitis and is thought to be a contributing factor in heart
 CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
 CC nodosum or pharyngitis. The polypeptides encoded by the open reading
 CC frames of the C. pneumoniae genome (see Y34584-Y35879) can be used in
 CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
 CC nucleotide sequences can also be used as immunogenic compositions,
 CC especially where the vector directs the expression of a neutralising
 CC epitope of C. pneumoniae.

SQ Sequence 621 AA;

Query Match 76.1%; Score 35; DB 20; Length 621;
 Best Local Similarity 87.5%; Pred. No. 38;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRIGHLYI 8
 |||||:
 DB 26 VRighlyl 33

RESULT 7

R25692
 ID R25692 standard; Protein; 189 AA.

XX

AC R25692;

DT 20-JAN-1993 (first entry)

DE Exons XV to XVIII of human hepatocyte growth factor.

KW HGF; enhance growth; preparing transgenic animals; hepatic disease;
 KW clinical diagnostic reagent; drug.

OS Homo sapiens.

PH Key Location/Qualifiers

FT Region 1..47

FT Region /label= exon XV

FT Region 48..82

FT Region /label= exon XVI

FT Region 83..131

FT Region /label= exon XVII

FT Region 132..187

FT Region /label= exon XVIII

PN JP04183394-A.

PD 30-JUN-1992.

PF 19-NOV-1990; 90JP-0314548.

PR 19-NOV-1990; 90JP-0314548.

PA (NAKA/) NAKAMURA T.

PA (TOYM) TOYONO KK.

DR WPI; 1992-265591/32.

DR N-PSDB; Q26727.

PT Recombinant human hepatocyte growth factor and DNA encoding it -

PT useful for diagnosis and treatment of hepatic disease and

PT transgenic animal prepn.

PS Disclosure; Page 22; 28pp; Japanese.

CC This sequence is made up of exons XV to XVIII of human hepatocyte

CC growth factor.

CC See also R25676-92, Q26713-27.

XX Sequence 189 AA;

SQ

Query Match 73.9%; Score 34; DB 13; Length 189;

Best Local Similarity 55.6%; Pred. No. 17;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VRIGHLYI 9
 :||:||||

DB 90 lrvahlyim 98

RESULT 8

Y98485
 ID Y98485 standard; Protein; 697 AA.

XX Y98485;

DT 31-JUL-2000 (first entry)

DE Hepatocyte growth factor Pep 20 used in nucleic acid transporter system.

KW Transporter system; nucleic acid delivery; gene therapy; cancer;

KW carcinogenesis; cardiovascular disease; infection.

OS Synthetic.

PN US6033884-A.


```

XX 07-MAR-2000.
PD
XX 14-DEC-1993; 930S-0167641.
PF
XX 20-MAR-1992; 920S-0855389.
PR 19-MAR-1993; 93WO-US02725.
XX
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
PI Gottchalk S, Sparrow J, Cristiano RJ, Woo SLC, Smith LC;
DR WPI; 2000-281993/24.
XX
PT System for transporting nucleic acid into cells; useful e.g. in gene
PT therapy and for generating transgenic animals; comprises binding agent
PT linked to nucleic acid, surface ligand and lytic agent -
XX
PS Disclosure; Figure 23A; 108pp; English.
XX
CC The present invention relates to a transporter system for delivering
CC nucleic acid to a cell. The system comprises a nucleic acid binding
CC complex, consisting of a binding molecule bonded non-covalently to the
CC nucleic acid, and covalently to a surface ligand, and a lytic agent. The
CC binding molecule is spermine or a spermidine derivative. Nucleotide
CC sequences A36633-A36652 and peptide sequences Y98456-Y98500 are used in
CC the construction of the transporter system of the invention. The
CC transporter system is used in gene therapy, particularly to deliver
CC nucleic acids to hepatocytes, muscle cells or bone forming cells, e.g for
CC treating cardiovascular disease, cancer, and infection. The transporter
CC systems are also used to create transgenic animals (as models for human
CC carcinogenesis or disease or for drug testing). Other uses include
CC transforming cells to produce proteins, or transfecting cells in vitro
CC to study the function of the nucleic acid. The use of a surface ligand
CC allows specific targeting of selected cells and tissues. The lytic agent
CC provides for release of the nucleic acid into the cellular interior, from
CC endosomes, without requiring endosomal or lysosomal degradation.
CC
XX
SQ Sequence 697 AA;

Query Match 73.9%; Score 34; DB 21; Length 697;
Best Local Similarity 55.6%; Pred. No. 70;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VRIGHLYIL 9
Db 598 lrvahlylm 606

RESULT 9
Y59030
ID Y59030 standard; peptide: 697 AA.
XX
AC Y59030;
XX
DT 07-MAR-2000 (first entry)
XX
DE Sequence of a peptide ligand Pep20.
XX
KW Nucleic acid transport system; NTS; cell surface receptor; cytosols;
KW nuclear membrane; lysis moiety; transgenic animal; human disease;
KW nucleic acid delivery; cancer.
XX
OS Synthetic.
XX
PN US5994109-A.
XX
PD 30-NOV-1999.
XX
PF 03-JUN-1995; 950S-0460890.
XX
PR 14-DEC-1993; 930S-0167641.

```

```

PR 20-MAR-1992; 920S-0855389.
PR 19-MAR-1993; 93WO-US02725.
PR 14-DEC-1993; 930S-0167641.
XX
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
PI Woo SLC, Cristiano RJ, Gottchalk S, Sparrow J, Smith LC;
DR WPI; 2000-038262/03.
XX
XX
PT Nucleic acid transport system, useful for creating transgenic animals
PT for assessing human disease such as cancer in an animal model -
XX
PS Disclosure; Fig 23A; 107pp; English.
XX
CC The invention relates to a nucleic acid transport system (NTS) for
CC delivering nucleic acid into a cell. The NTS contains but is not limited
CC to 5 components: (a) the nucleic acid or a macromolecule to be delivered;
CC (b) a moiety that recognizes and binds to a cell surface receptor or
CC antigen or is capable of entering a cell through cytosols; (c) a nucleic
CC acid or macromolecular molecule binding moiety; (d) a moiety that is
CC capable of moving or initiating movement through a nuclear membrane; and/
CC or (e) a lysis moiety that enables the transport of the entire complex
CC from the cell surface directly into the cytoplasm of the cell. The NTS
CC delivers nucleic acid into the cellular interior as well as the nucleus
CC of specific cells. The NTS can be used to treat disorders by targeting
CC specific nucleic acid accordingly. The NTS can also be used to create
CC transgenic animals for assessing human disease, such as cancer, in an
CC animal model. The NTS can be used in vitro with tissue culture cells
CC which allows the role of various nucleic acids to be studied by targeting
CC specific expression into specifically targeted tissue culture cells. The
CC lysis agent within the NTS avoids the problem of endosomal/lysosomal
CC degradation.
CC
XX
SQ Sequence 697 AA;

Query Match 73.9%; Score 34; DB 21; Length 697;
Best Local Similarity 55.6%; Pred. No. 70;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VRIGHLYIL 9
Db 598 lrvahlylm 606

RESULT 10
B45838
ID B45838 standard; Protein; 697 AA.
XX
AC B45838;
XX
DT 21-MAR-2001 (first entry)
XX
DE Nucleic acid transporter system peptide ligand SEQ ID NO 50.
XX
KW Nucleic acid delivery; nucleic acid transporter system; hormone; enzyme;
KW growth factor; clotting factor; apolipoprotein; receptor; drug; oncogene;
KW tumor antigen; tumor suppressor; viral antigen; parasitic antigen;
KW bacterial antigen.
XX
OS Unidentified.
XX
PN US6150168-A.
XX
PD 21-NOV-2000.
XX
PF 05-JUN-1995; 950S-0460971.
XX
PR 14-DEC-1993; 930S-0167641.
PR 20-MAR-1992; 920S-0855389.
PR 19-MAR-1993; 93WO-US02725.
XX

```

PA (BAYU) BAYLOR COLLEGE MEDICINE.
 XX Gottchalk S, Sparrow J, Cristiano RJ, Smith LC, Woo SLG;
 XX WPI; 2001-049093/06.
 DR Nucleic acid transporter system for delivering nucleic acid into a
 XX cell, useful for delivering proteins and polypeptides to cells,
 PT including growth factors, enzymes, hormones, and tumor suppressors -
 XX
 PS Claim 8; Column 115-118; 105pp; English.
 XX
 CC This invention describes a novel system (I) for delivering a nucleic acid
 CC to a cell, comprising a binding complex comprising a ligand binding
 CC molecule noncovalently bound to a nucleic acid and covalently linked to a
 CC surface ligand, and a second binding complex comprising a second binding
 CC molecule noncovalently bound to a nucleic acid and covalently linked to a
 CC nuclear ligand. The complexes are simultaneously bound to the nucleic
 CC acid. The nucleic acid transporter system can also be used in a method
 CC for the in vivo targeting of the insertion of DNA into a cell. It can
 CC also be used in processes for producing transformed cell lines. The
 CC system can be used to deliver a variety of proteins and polypeptides,
 CC such as hormones, growth factors, enzymes, clotting factors,
 CC apolipoproteins, receptors, drugs, oncogenes, tumor antigens, tumor
 CC suppressors, viral antigens, parasitic antigens, and bacterial antigens.
 CC The transporter system uses lysis agents to overcome the problems of
 CC endosomal/lysosomal degradation seen with prior art systems.
 XX
 SQ Sequence 697 AA;
 XX
 Query Match 73.9%; Score 34; DB 22; Length 697;
 Best Local Similarity 55.6%; Pred. No. 70;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 OY 1 VRIGHLYTL 9
 :|:|||||
 DB 598 lrvahlylm 606
 XX
 RESULT 11
 R07144
 ID R07144 standard; protein; 723 AA.
 XX
 AC R07144;
 XX
 DT 31-JAN-1991 (first entry)
 XX
 DE Tumour cytotoxic factor II.
 XX
 KW TCF II; glycoprotein; fibroblasts; anticancer agent.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..29
 XX
 PN WO9010651-A.
 XX
 PD 20-SEP-1990.
 XX
 PF 03-SEP-1990; 90WO-J000314.
 XX
 PR 16-JAN-1990; 90JP-0006692.
 PR 03-OCT-1989; 89JP-0058631.
 XX
 PA (SNOW) SNOW BRAND MILK PRODUCTS.
 XX
 PI Higashio K; Mistuda S; Shima N; Itagaki Y; Nagao M;
 DR WPI; 1990-304990/40.
 DR P-PSDB; R07144.
 XX

PT Human glycoprotein having antitumor activity - for inducing
 PT differentiation of leukemia cells and enhancing cell mediated
 PT immunity.
 XX
 PS Disclosure; Fig 15; 73pp; Japanese.
 XX
 CC The sequence was deduced from the cDNA obtd. from a clone isolated
 CC from a cDNA library prepd. from RNA extd. from fibroblast IMR-90
 CC cells (ATCC CCL-186). The N-terminal of the alpha chain is
 CC unclear. The sequence shows homology to hGF except between AAs
 CC 162 and 166 of this sequence which is absent from the TCF-II
 CC sequence. The DNA can be used to produce the TCF II by recombinant
 CC DNA technology. TCF II is an anticancer agent and induces
 CC differentiation of leukemia cells, enhances cell-mediated immunity,
 CC and accelerates the proliferation of human blood vessel endothelial
 CC cells and hepatic parenchymal cells.
 XX
 SQ Sequence 723 AA;
 XX
 Query Match 73.9%; Score 34; DB 11; Length 723;
 Best Local Similarity 55.6%; Pred. No. 73;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 OY 1 VRIGHLYTL 9
 :|:|||||
 DB 624 lrvahlylm 632
 XX
 RESULT 12
 R14307
 ID R14307 standard; Protein; 723 AA.
 XX
 AC R14307;
 XX
 DT 10-JAN-1992 (first entry)
 XX
 DE Plasminogen-like growth factor protein.
 XX
 KW Mitogenic activity; PLGF.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..32
 FT /tag- a
 FT /note- "signal peptide"
 FT 33..723
 FT /tag- b
 FT /note- "mature peptide"
 FT 489..490
 FT /tag- c
 FT /note- "conserved, relative to plasminogen"
 FT 529
 FT Active-site /tag- d
 XX
 PN US7582063-A.
 XX
 PD 17-SEP-1991.
 XX
 PF 14-SEP-1990; 90US-0582063.
 XX
 PR 14-SEP-1990; 90US-0582063.
 XX
 PA (USSH) NAT INST OF HEALTH.
 XX
 PI Rubin JS, Chan AML, Aaronson SA;
 XX
 DR WPI; 1991-317957/43.
 DR N-PSDB; Q14182.
 XX
 PT New plasminogen-like growth factor protein - having potent
 PT mitogenic activity on melanocytes, epithelial cells and

PT endothelial cells
XX
PS Disclosure; Fig 6; 33pp; English.
XX

CC The amino acid sequence is that of a plasminogen-like growth
CC factor (PEGF) protein. It can be isolated from conditioned medium
CC of human embryonic lung fibroblasts. The DNA encoding PEGF can be
CC obtd. from a DNA library of M246 human embryonic lung fibroblasts.
CC The sequence may be used in the determ. of the level of mRNA
CC transcripts encoding the protein which may be useful in the
CC diagnosis of malignancy. The amino acid sequence shown here is
CC that given in the specification, it differs from the expected
CC translation of the PEGF gene at amino acids 175 - pro, which
CC translates as Thr from the given gene.
XX
SQ Sequence 723 AA;

Query Match 73.9%; Score 34; DB 12; Length 723;
Best Local Similarity 55.6%; Pred. No. 73;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VRIGHLYIL 9
: : : : :
DB 624 Irvahlyim 632

RESULT 13
R15624
ID R15624 standard; Protein; 723 AA.
XX
AC R15624;

DT 18-MAR-1992 (first entry)
XX
DE Human leukocyte-derived HGF encoded by clone HLC3.
XX
KW Hepatocyte growth factor; liver; hepatoma.
XX
OS Homo sapiens.

PN EP461560-A.
XX
PD 18-DEC-1991.

PF 07-JUN-1991; 91EP-0109369.
XX
PR 11-JUN-1990; 90JP-0152474.

PA (TOYOM) TOYO BOSEKI KK.

PI Nakamura T, Hagiya M, Seki T, Shimonishi M, Shimizu S;
PI Ihara T, Sakaguchi M, Asami O;

DR WPI; 1991-370578/51.
XX
N-PSDB; Q15177.

PT Recombinant human leukocyte-derived hepatocyte growth factor -
PT With DNA encoding it, recombinant expression vectors and
PT transforment cells expressing it.
XX
PS Claim 3; Fig 3; 33pp; English.

XX The sequence was deduced from a portion of HLC3, one of two clones,
CC (for HLC2 see R15623) isolated from a cDNA library prepd. from mRNA
CC extracted from human leukocytes. HLC2 has almost the same sequence
CC as HLC3 except for five residues (162-166) in HLC2 which do not
CC appear in HLC3. HLC3 shows similar characteristics to the human
CC liver-derived HGF identified in Nature, 342, 440, 1989, but differs
CC at 14 positions in the amino acid sequence. The DNA sequence can be
CC expressed and the resulting protein, recombinant HGF, used in hepa-
CC tocyte cultivation, liver regeneration, hepatocyte research, esp.
CC into the mechanism of hepatoma, and to prepare anti-HGF antibodies

CC for diagnosis and therapy.
XX
SQ Sequence 723 AA;

Query Match 73.9%; Score 34; DB 12; Length 723;
Best Local Similarity 55.6%; Pred. No. 73;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VRIGHLYIL 9
: : : : :
DB 624 Irvahlyim 632

RESULT 14
R21142
ID R21142 standard; Protein; 723 AA.
XX
AC R21142;

DT 20-MAY-1992 (first entry)
XX
DE Human TCF-II.

KW Tumour cell killing factor; hepatocyte growth factor.
XX
OS Homo sapiens.

FT Key Location/Qualifiers
FT Region 490..505
/note="Beta-chain N-terminal sequence"

PN W09201053-A.
XX
PD 23-JAN-1992.

PF 15-JUL-1991; 91WO-JP00942.

PR 13-JUL-1990; 90JP-0185852.

PA (SNOW) SNOW BRAND MILK PRODUCTS.

PI Shima N, Higashio K, Nagao M, Oogaki F, Takaoka H;
PI Tsuda E;

DR WPI; 1992-056868/07.
XX
N-PSDB; Q21066.

PT Prodn. of liver cell growth factor TCF-II - by culture of a
PT transforment contg. cDNA for TCF-II and derivling from human
PT foetal lung fibroblast cells
XX
PS Claim 6; Fig 1; 47pp; Japanese.

XX A human foetal lung fibroblast library was screened with a synthetic
CC probe coding for the N-terminal sequence of TCF-II. A cDNA clone was
CC isolated and sequenced. The coding sequence can be incorporated into
CC a plasmid and transformed into a microorganism for production of TCF-
CC II. The protein itself can be used as a liver cell growth factor or a
CC tumour cell killing factor. See also Q21067-9.
XX
SQ Sequence 723 AA;

Query Match 73.9%; Score 34; DB 13; Length 723;
Best Local Similarity 55.6%; Pred. No. 73;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VRIGHLYIL 9
: : : : :
DB 624 Irvahlyim 632

RESULT 15

R25677 ID R25677 standard; Protein; 723 AA.

XX AC R25677;

XX DT 20-JAN-1993 (first entry)

XX DE Recombinant human hepatocyte growth factor.

XX KW HGF; enhance growth; preparing transgenic animals; hepatic disease;

XX KM clinical diagnostic reagent; drug.

XX OS Homo sapiens.

XX PN JP04183394-A.

XX PD 30-JUN-1992.

XX PF 19-NOV-1990; 90JP-0314548.

XX PR 19-NOV-1990; 90JP-0314548.

XX PA (NAKA/) NAKAMURA T.

XX PA (TOYM) TOYOBO KK.

XX DR WPI; 1992-265591/32.

XX PT Recombinant human hepatocyte growth factor and DNA encoding it -
 PT useful for diagnosis and treatment of hepatic disease and
 PT transgenic animal prepn.

XX PS Disclosure; Page 13; 28pp; Japanese.

XX CC This sequence represents a recombinant human hepatocyte growth
 CC factor. It has physiological activity, and using it enhanced growth
 CC of hepatocytes is possible. It is useful as a clinical diagnostic
 CC reagent, or a drug for treating hepatic disease.
 CC See also R25676-92, Q26713-27.

XX SQ Sequence 723 AA;

Query Match 73.9%; Score 34; DB 13; Length 723;

Best Local Similarity 55.6%; Pred. No. 73;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VRIGHLYIL 9

DB 624 lrvahlyim 632

Search completed: June 20, 2001, 14:04:01
 Job time: 48 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2001, 14:09:31 ; Search time 16.63 Seconds
(without alignments)
20.599 Million cell updates/sec

Title: US-09-692-401-5
Perfect score: 50
Sequence: 1 VVRIGHLYIL 10

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	100.0	314	1	MAGC_HUMAN
2	37	74.0	314	1	MAG2_HUMAN
3	36	72.0	421	1	CD31_ARATH
4	35	70.0	613	1	CCPE_CHPN
5	35	70.0	692	1	GTRE_BABA
6	35	70.0	728	1	HGF_HUMAN
7	35	70.0	728	1	HGF_MOUSE
8	35	70.0	728	1	HGF_RAT
9	35	70.0	933	1	AT1D_HUMAN
10	34	68.0	424	1	CD31_SOLFU
11	33	66.0	314	1	MAG3_HUMAN
12	33	66.0	538	1	TNSE_ECOLI
13	33	66.0	607	1	EMPD_PIG
14	33	66.0	619	1	GCKR_XENLA
15	32	64.0	190	1	SOMA_CROMO
16	32	64.0	256	1	TIPA_PHAVU
17	32	64.0	368	1	SNCI_HUMAN
18	32	64.0	404	1	SGAA_HYPER
19	32	64.0	415	1	ROF_HUMAN
20	32	64.0	449	1	ROH1_HUMAN
21	32	64.0	449	1	ROH2_HUMAN
22	32	64.0	510	1	C312_DROME
23	32	64.0	617	1	ETFD_HUMAN
24	32	64.0	1374	1	RPOB_RICPR
25	32	64.0	1376	1	RPOB_RICPR
26	32	64.0	2241	1	TEGU_HICWA
27	31	62.0	213	1	YLL1_ARCFU
28	31	62.0	358	1	LEU3_THIPE
29	31	62.0	425	1	SYH_METH
30	31	62.0	484	1	MURC_BUCHI
31	31	62.0	504	1	C6H5_DROME
32	31	62.0	576	1	STPC_YEAST
33	31	62.0	599	1	PAKE_BORBO

34	31	62.0	619	1	YOL8_CAEEL	Q02335 caenorhabdi
35	31	62.0	658	1	CPT2_MOUSE	P52825 mus musculu
36	31	62.0	658	1	CPT2_MOUSE	P18886 rattus norv
37	31	62.0	775	1	PEL_STRMU	O59334 streptococc
38	30	60.0	206	1	UPP_MYCE	P47276 mycoplasma
39	30	60.0	206	1	UPP_MYCN	P75081 mycoplasma
40	30	60.0	213	1	ATPO_BOVIN	P13621 bos taurus
41	30	60.0	213	1	ATPO_HUMAN	P48047 homo sapien
42	30	60.0	213	1	ATPO_RAT	Q06647 rattus norv
43	30	60.0	229	1	CREB_ECOLI	P08368 escherichia
44	30	60.0	238	1	DEOD_HAEIN	P44417 haemophilus
45	30	60.0	288	1	Y077_HAEIN	P43935 haemophilus

ALIGNMENTS

```

RESULT 1
MAGC_HUMAN STANDARD; PRT; 314 AA.
AC P43365;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE MELANOMA-ASSOCIATED ANTIGEN 12 (MAGE-12 ANTIGEN) (MAGE12F).
GN MAGE12 OR MAGE12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94102805; PubMed=8276455;
de Smet C., Lurguin C., van der Bruggen P., de Plaen E., Brasseur F.,
RA Boon T.;
RT "Sequence and expression pattern of the human MAGE2 gene.";
RL Immunogenetics 39:121-129(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94311935; PubMed=8037761;
RL Tissue-Skin;
RA Ding M., Beck R.J., Keller C.J., Fenton R.G.;
RT "Cloning and analysis of MAGE-1-related genes.";
RL Biochem. Biophys. Res. Commun. 202:549-555(1994).
CC -1- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN TUMOR
CC TRANSFORMATION OR ASPECTS OF TUMOR PROGRESSION.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES.
CC -1- SIMILARITY: BELONGS TO THE MAGE FAMILY. STRONG SIMILARITY (90%)
CC WITH MAGE-2.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: L18877; AAA19023.1; -.
CC MIM: 300177; -.
CC InterPro: IPR002190; -.
CC Pfam: PF01454; MAGE; 1.
CC Antigen: Multigene family; Tumor antigen.
CC DOMAIN 40 43 POLY-SER.
CC CONFLICT 10 10 C -> S (IN REF. 2).
CC CONFLICT 187 187 A -> D (IN REF. 2).
CC CONFLICT 300 300 P -> S (IN REF. 2).
CC SEQUENCE 314 AA; 34802 MW; 3E0787CECD8816A5 CRC64;

```

Query Match 100.0%; Score 50; DB 1; Length 314;
 Best Local Similarity 100.0%; Pred. No. 0.01;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVRIGHLYIL 10
 |||||
 Db 169 VVRIGHLYIL 178

RESULT 2
 MAG2_HUMAN STANDARD; PRT; 314 AA.
 AC P43356;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE MELANOMA-ASSOCIATED ANTIGEN 2 (MAGE-2 ANTIGEN).
 GN MAGE2 OR MAGE2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94102805; PubMed=8276455;
 RA de Smet C., Luquin C., van der Bruggen P., de Plaen E., Brasseur F.,
 RA Boon T.;
 RT "Sequence and expression pattern of the human MAGE2 gene.";
 RL Immunogenetics 39:121-129(1994).
 RN [2]
 RP MUYAGENESIS.
 RC TISSUE-BLOOD;
 RX MEDLINE=94157413; PubMed=8113684;
 RA Gaugier B., van den Eynde B., van der Bruggen P., Romero P.,
 RA Gattor J.J., de Plaen E., Lethe B., Brasseur F., Boon T.;
 RT "Human gene MAGE-3 codes for an antigen recognized on a melanoma by
 RT autologous cytolytic T lymphocytes.";
 RL J. Exp. Med. 179:921-930(1994).
 CC -1- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
 CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
 CC PROGRESSION. ANTIGEN RECOGNIZED ON A MELANOMA BY AUTOLOGOUS
 CC CYTOLYTIC T LYMPHOCYTES.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
 CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG
 CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
 CC FOR TESTES. BELONGS TO THE MAGE FAMILY. STRONG SIMILARITY (90%)
 CC WITH MAGE-12.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L18920; AAA17729.1; -
 CC DR MIM: 300173; -
 CC DR InterPro: IPR002190; -
 CC DR Pfam: PF01454; MAGE: 1.
 KW Antigen; Multigene family; Tumor antigen.
 FT DOMAIN 40 43 POLY-SER.
 FT MUTAGEN 170 170 V->D: IMPROVES ABILITY TO BIND TO HLA-A1.
 SO SEQUENCE 314 AA; 35055 MW; 844F16335A2BEC7 CRC64;

Db 169 VVRISHLYIL 178

RESULT 3
 CDSL_ARATH STANDARD; PRT; 421 AA.
 AC O04928; O48808;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PHOSPHATIDATE CYTIDYLTRANSFERASE (EC 2.7.7.41) (CDP-DIGLYCERIDE
 DE SYNTHETASE) (CDP-DIGLYCERIDE PYROPHOSPHORYLASE) (CDP-DIACYLGlycerol
 DE SYNTHASE) (CDP-DG SYNTHETASE).
 GN CDSL OR F2401.17.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SYRAIN-CV. C24; TISSUE=Flower, and Siliques;
 RX MEDLINE=97239925; PubMed=9085581;
 RA Kopke J., Ludwig M., Mueller-Roeber B.;
 RT "Complementary DNAs encoding eukaryotic-type cytidine-5'-diphosphate-
 RT diacylglycerol synthases of two plant species.";
 RL Plant Physiol. 113:997-1002(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Shin P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altafi H., Bel B., Chin C., Chio J., Choi E.,
 RA Con L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu J., Liu S., Mukharsy N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thayer J.A., Tortini M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.;
 RL Submitted (Mar 2000) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: MAY BE INVOLVED IN THE SYNTHESIS OF MINOR PHOSPHOLIPIDS
 CC -1- AND IN MODULATION OF IP3-MEDIATED SIGNAL TRANSDUCTION.
 CC -1- CATALYTIC ACTIVITY: CTP + PHOSPHATIDATE = PYROPHOSPHATE +
 CC CDP-DIACYLGlycerol.
 CC -1- PATHWAY: PHOSPHOLIPID BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE CDS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X94306; CA63969.1; -
 CC DR EMBL: AC003113; AAF70845.1; -
 CC DR Mendel: 27238; Arath:2834;27238.
 CC DR InterPro: IPR000374; -
 CC DR Pfam: PF01148; Cytidylyltrans: 1.
 CC DR PROSITE: PS01315; CDS: 1.
 KW Transferase; Nucleotidyltransferase; Phospholipid biosynthesis;
 KW Transmembrane
 FT TRANSMEM 60 80 POTENTIAL.
 FT TRANSMEM 149 169 POTENTIAL.
 FT TRANSMEM 206 226 POTENTIAL.
 FT TRANSMEM 246 266 POTENTIAL.
 FT TRANSMEM 321 341 POTENTIAL.
 SO SEQUENCE 421 AA; 48659 MW; 96D238DB4B913AA3 CRC64;

Query Match 74.0%; Score 37; DB 1; Length 314;
 Best Local Similarity 80.0%; Pred. No. 4;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Query Match 72.0%; Score 36; DB 1; Length 421;
 Best Local Similarity 77.8%; Pred. No. 8.6;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VVRIGHLYI 9
11:11111
Db 66 VVRIGHLYI 74

RESULT 4
GCFE_CHLPN STANDARD; PRT; 613 AA.
ID GCFE_CHLPN
AC G928H0; G9J095;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GCFE PROTEIN HOMOLOG.
GN GCFE OR CPN0373 OR CP0383.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kallman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
"Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RT Nat. Genet. 21:385-389(1999).
RL 12
RN SEQUENCE FROM N.A.
RP STRAIN-AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
White O., Hickey E.K., Peterson J., Utterback T., Betty K., Bass S.,
Linher K., Weisman J., Khouri H., Craven B., Bowman C., Dodson R.,
Eisen J., Fraser C.W.,
"Genome sequences of Chlamydia trachomatis Moyn and Chlamydia
pneumoniae AR39.";
RT Nucleic Acids Res. 28:1397-1406(2000).
RL 13
RN SEQUENCE FROM N.A.
RP STRAIN-J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tsubuchi M., Kishi F., Ouchi K.,
Shiba T., Ishii K., Hattori M., Kohara S., Nakazawa F.,
"Comparison of whole genome sequences of Chlamydia pneumoniae J138
from Japan and CWL029 from USA.";
RT Nucleic Acids Res. 28:2311-2314(2000).
RL 1- SIMILARITY: BELONGS TO THE GCFE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001621; AAD18517.1; -
DR EMBL: AE002200; AAF38230.1; -
DR EMBL: AP002546; BAA9581.1; -
DR TIGR: CP0383;
SO SEQUENCE 613 AA; 68308 MW; 40B1F831E02DC6D4 CRC64;

Query Match 70.0%; Score 35; DB 1; Length 613;
Best Local Similarity 87.5%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 VVRIGHLYI 9
111111111
Db 18 VVRIGHLYI 25

RESULT 5
GYR_BABA STANDARD; PRT; 692 AA.
ID GYR_BABA
AC P94281;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DNA GYRASE SUBUNIT B (EC 5.99.1.3).
GN GYR_B.
OS Bartonella bacilliformis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bartonellaceae; Bartonella.
OX NCBI_TaxID=774;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-KC383 / ATCC 35685;
RX MEDLINE=99013570; PubMed=9797224;
RA Battisti J.M., Smitherman L.S., Samuels D.S., Minnick M.F.,
"Mutations in Bartonella bacilliformis gyrB confer resistance to
RT coumermycin A1.";
RL Antimicrob. Agents Chemother. 42:2906-2913(1998).
CC -1- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-
CC STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE
CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
CC -1- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING
CC OF DOUBLE-STRANDED DNA.
CC -1- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
CC BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
CC ENZYME FORMS AN A2B2 TETRAMER.
CC -1- SIMILARITY: BELONGS TO TYPE II TOPOISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U82225; AAC71079.1; -
DR InterPro: IPR001241; -
DR InterPro: IPR002288; -
DR Pfam: PF00986; DNA_gyraseB_C; 1.
DR Pfam: PF00204; DNA_topoisom; 1.
DR PRINTS: PR00418; TP12FAMILY.
DR PROSITE: PS00117; TOPOISOMERASE_II; 1.
KW Topoisomerase; Isomerase; ATP-binding; Antibiotic resistance.
FT VARIANT 124 124 G -> S (IN COMMERMYCIN A1 MUTANT).
FT VARIANT 184 184 R -> Q (IN COMMERMYCIN A1 MUTANT).
FT VARIANT 214 214 T -> A (IN COMMERMYCIN A1 MUTANT).
FT VARIANT 214 214 T -> I (IN COMMERMYCIN A1 MUTANT).
SO SEQUENCE 692 AA; 77445 MW; F54948EEF342281F CRC64;

Query Match 70.0%; Score 35; DB 1; Length 692;
Best Local Similarity 55.6%; Pred. No. 23;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VVRIGHLYI 9
111111111
Db 575 VVRIGHLYI 593

RESULT 6
HGF_HUMAN STANDARD; PRT; 728 AA.
ID HGF_HUMAN
AC P14210;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HEPATOCYTE GROWTH FACTOR PRECURSOR (SCATTER FACTOR) (SFG
(HEPATOPOIETIN A)).

GN HGF OR HPTA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91340155; PubMed=1831432;
 RA Seki T., Hagiya M., Shimomishi M., Nakamura T., Shimizu S.;
 RT "Organization of the human hepatocyte growth factor-encoding gene";
 RL Gene 102:213-219(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=89392017; PubMed=2528952;
 RA Miyazawa K., Tsubouchi H., Naka D., Takahashi K., Okigaki M.,
 RA Arakaki N., Nakayama H., Hirono S., Sakiyama O., Takahashi K.,
 RA Gonda E., Daikuhara Y., Kitamura N.;
 RT "Molecular cloning and sequence analysis of cDNA for human hepatocyte
 RT growth factor";
 RL Biochem. Biophys. Res. Commun. 163:967-973(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leukocyte;
 RX MEDLINE=91025062; PubMed=2145836;
 RA Seki T., Ihara I., Sugimura A., Shimomishi M., Nishizawa T.,
 RA Asami O., Hagiya M., Nakamura T., Shimizu S.;
 RT "Isolation and expression of cDNA for different forms of hepatocyte
 RT growth factor from human leukocyte";
 RL Biochem. Biophys. Res. Commun. 172:321-327(1990).
 RN [4]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 55-73 AND 495-520.
 RC TISSUE=Liver;
 RX MEDLINE=90066676; PubMed=2531289;
 RA Nakamura T., Nishizawa T., Hagiya M., Seki T., Shimomishi M.,
 RA Sugimura A., Tashiro K., Shimizu S.;
 RT "Molecular cloning and expression of human hepatocyte growth factor";
 RL Nature 342:440-443(1989).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryonic fibroblast;
 RX MEDLINE=91334393; PubMed=1831266;
 RA Weidner K.M., Arakaki N., Hartmann G., Vandekerckhove J., Weingart S.,
 RA Rieder H., Fomatsch C., Tsubouchi H., Hishida T., Daikuhara Y.,
 RA Birnmeier W.;
 RT "Evidence for the identity of human scatter factor and human
 RT hepatocyte growth factor";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:7001-7005(1991).
 RN [6]
 RP SIGNAL SEQUENCE CLEAVAGE SITE.
 RX MEDLINE=91207365; PubMed=1826837;
 RA Yoshiyama Y., Arakaki N., Naka D., Takahashi K., Hirono S., Kondo J.,
 RA Nakayama H., Gonda E., Kitamura N., Tsubouchi H., Ishi T.,
 RA Hishida T., Daikuhara Y.;
 RT "Identification of the N-terminal residue of the heavy chain of both
 RT native and recombinant human hepatocyte growth factor";
 RL Biochem. Biophys. Res. Commun. 175:660-667(1991).
 RN [7]
 RP CARBOHYDRATE-LINKAGE SITE 476.
 RX MEDLINE=9129192; PubMed=1482348;
 RA Shimizu N., Hara H., Sogabe T., Sakai H., Ihara I., Inoue H.,
 RA Nakamura T., Shimizu S.;
 RT "Hepatocyte growth factor is linked by O-glycosylated oligosaccharide
 RT on the alpha chain";
 RL Biochem. Biophys. Res. Commun. 189:1329-1335(1992).
 RN [8]
 RP MUTAGENESIS.
 RX MEDLINE=92331602; PubMed=1321034;
 RA Lokker N.A., Mark M.R., Luis E.A., Bennett G.L., Robbins K.A.,
 RA Baker J.B., Godowski P.J.;
 RT "Structure-function analysis of hepatocyte growth factor:
 RT identification of variants that lack mitogenic activity yet retain
 RT high affinity receptor binding";

RL EMBO J. 11:2503-2510(1992).
 RN [9]
 RP STRUCTURE BY NMR OF 31-127.
 RX MEDLINE=98154323; PubMed=9493272;
 RA Zhou H., Mazulla M.J., Kaufman J.D., Stahl S.J., Wingfield P.T.,
 RA Rubin J.S., Bottaro D.P., Byrd R.A.;
 RT "The solution structure of the N-terminal domain of hepatocyte growth
 RT factor reveals a potential heparin-binding site";
 RL Structure 6:109-116(1998).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 35-210.
 RX MEDLINE=99036858; PubMed=9817840;
 RA Ullsch M., Lokker N.A., Godowski P.J., de Vos A.M.;
 RT "Crystal structure of the NK1 fragment of human hepatocyte growth
 RT factor at 2.0-A resolution";
 RL Structure 6:1383-1393(1998).
 CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
 CC HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPIC FACTOR, AND ACTS
 CC AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
 CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.
 CC -1- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
 CC DISULFIDE BOND.
 CC -1- SIMILARITY: CONTAINS 4 KRINGLE REGIONS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY, PLASMINOGEN SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D90334; BAA14348.1; -;
 DR EMBL: D90334; BAA14348.1; JOINED.
 DR EMBL: D90339; BAA14348.1; JOINED.
 DR EMBL: D90320; BAA14348.1; JOINED.
 DR EMBL: D90322; BAA14348.1; JOINED.
 DR EMBL: D90323; BAA14348.1; JOINED.
 DR EMBL: D90324; BAA14348.1; JOINED.
 DR EMBL: D90325; BAA14348.1; JOINED.
 DR EMBL: D90326; BAA14348.1; JOINED.
 DR EMBL: D90327; BAA14348.1; JOINED.
 DR EMBL: D90328; BAA14348.1; JOINED.
 DR EMBL: D90329; BAA14348.1; JOINED.
 DR EMBL: D90330; BAA14348.1; JOINED.
 DR EMBL: D90331; BAA14348.1; JOINED.
 DR EMBL: D90332; BAA14348.1; JOINED.
 DR EMBL: D90333; BAA14348.1; JOINED.
 DR EMBL: D90334; BAA14348.1; JOINED.
 DR EMBL: D90335; BAA14348.1; JOINED.
 DR EMBL: D90336; BAA14348.1; JOINED.
 DR EMBL: D90337; BAA14348.1; JOINED.
 DR EMBL: D90338; BAA14348.1; JOINED.
 DR EMBL: D90339; BAA14348.1; JOINED.
 DR EMBL: D90340; BAA14348.1; JOINED.
 DR EMBL: D90341; BAA14348.1; JOINED.
 DR EMBL: D90342; BAA14348.1; JOINED.
 DR EMBL: D90343; BAA14348.1; JOINED.
 DR EMBL: D90344; BAA14348.1; JOINED.
 DR EMBL: D90345; BAA14348.1; JOINED.
 DR EMBL: D90346; BAA14348.1; JOINED.
 DR EMBL: D90347; BAA14348.1; JOINED.
 DR EMBL: D90348; BAA14348.1; JOINED.
 DR EMBL: D90349; BAA14348.1; JOINED.
 DR EMBL: D90350; BAA14348.1; JOINED.
 DR EMBL: D90351; BAA14348.1; JOINED.
 DR EMBL: D90352; BAA14348.1; JOINED.
 DR EMBL: D90353; BAA14348.1; JOINED.
 DR EMBL: D90354; BAA14348.1; JOINED.
 DR EMBL: D90355; BAA14348.1; JOINED.
 DR EMBL: D90356; BAA14348.1; JOINED.
 DR EMBL: D90357; BAA14348.1; JOINED.
 DR EMBL: D90358; BAA14348.1; JOINED.
 DR EMBL: D90359; BAA14348.1; JOINED.
 DR EMBL: D90360; BAA14348.1; JOINED.
 DR EMBL: D90361; BAA14348.1; JOINED.
 DR EMBL: D90362; BAA14348.1; JOINED.
 DR EMBL: D90363; BAA14348.1; JOINED.
 DR EMBL: D90364; BAA14348.1; JOINED.
 DR EMBL: D90365; BAA14348.1; JOINED.
 DR EMBL: D90366; BAA14348.1; JOINED.
 DR EMBL: D90367; BAA14348.1; JOINED.
 DR EMBL: D90368; BAA14348.1; JOINED.
 DR EMBL: D90369; BAA14348.1; JOINED.
 DR EMBL: D90370; BAA14348.1; JOINED.
 DR EMBL: D90371; BAA14348.1; JOINED.
 DR EMBL: D90372; BAA14348.1; JOINED.
 DR EMBL: D90373; BAA14348.1; JOINED.
 DR EMBL: D90374; BAA14348.1; JOINED.
 DR EMBL: D90375; BAA14348.1; JOINED.
 DR EMBL: D90376; BAA14348.1; JOINED.
 DR EMBL: D90377; BAA14348.1; JOINED.
 DR EMBL: D90378; BAA14348.1; JOINED.
 DR EMBL: D90379; BAA14348.1; JOINED.
 DR EMBL: D90380; BAA14348.1; JOINED.
 DR EMBL: D90381; BAA14348.1; JOINED.
 DR EMBL: D90382; BAA14348.1; JOINED.
 DR EMBL: D90383; BAA14348.1; JOINED.
 DR EMBL: D90384; BAA14348.1; JOINED.
 DR EMBL: D90385; BAA14348.1; JOINED.
 DR EMBL: D90386; BAA14348.1; JOINED.
 DR EMBL: D90387; BAA14348.1; JOINED.
 DR EMBL: D90388; BAA14348.1; JOINED.
 DR EMBL: D90389; BAA14348.1; JOINED.
 DR EMBL: D90390; BAA14348.1; JOINED.
 DR EMBL: D90391; BAA14348.1; JOINED.
 DR EMBL: D90392; BAA14348.1; JOINED.
 DR EMBL: D90393; BAA14348.1; JOINED.
 DR EMBL: D90394; BAA14348.1; JOINED.
 DR EMBL: D90395; BAA14348.1; JOINED.
 DR EMBL: D90396; BAA14348.1; JOINED.
 DR EMBL: D90397; BAA14348.1; JOINED.
 DR EMBL: D90398; BAA14348.1; JOINED.
 DR EMBL: D90399; BAA14348.1; JOINED.
 DR EMBL: D90400; BAA14348.1; JOINED.
 DR EMBL: D90401; BAA14348.1; JOINED.
 DR EMBL: D90402; BAA14348.1; JOINED.
 DR EMBL: D90403; BAA14348.1; JOINED.
 DR EMBL: D90404; BAA14348.1; JOINED.
 DR EMBL: D90405; BAA14348.1; JOINED.
 DR EMBL: D90406; BAA14348.1; JOINED.
 DR EMBL: D90407; BAA14348.1; JOINED.
 DR EMBL: D90408; BAA14348.1; JOINED.
 DR EMBL: D90409; BAA14348.1; JOINED.
 DR EMBL: D90410; BAA14348.1; JOINED.
 DR EMBL: D90411; BAA14348.1; JOINED.
 DR EMBL: D90412; BAA14348.1; JOINED.
 DR EMBL: D90413; BAA14348.1; JOINED.
 DR EMBL: D90414; BAA14348.1; JOINED.
 DR EMBL: D90415; BAA14348.1; JOINED.
 DR EMBL: D90416; BAA14348.1; JOINED.
 DR EMBL: D90417; BAA14348.1; JOINED.
 DR EMBL: D90418; BAA14348.1; JOINED.
 DR EMBL: D90419; BAA14348.1; JOINED.
 DR EMBL: D90420; BAA14348.1; JOINED.
 DR EMBL: D90421; BAA14348.1; JOINED.
 DR EMBL: D90422; BAA14348.1; JOINED.
 DR EMBL: D90423; BAA14348.1; JOINED.
 DR EMBL: D90424; BAA14348.1; JOINED.
 DR EMBL: D90425; BAA14348.1; JOINED.
 DR EMBL: D90426; BAA14348.1; JOINED.
 DR EMBL: D90427; BAA14348.1; JOINED.
 DR EMBL: D90428; BAA14348.1; JOINED.
 DR EMBL: D90429; BAA14348.1; JOINED.
 DR EMBL: D90430; BAA14348.1; JOINED.
 DR EMBL: D90431; BAA14348.1; JOINED.
 DR EMBL: D90432; BAA14348.1; JOINED.
 DR EMBL: D90433; BAA14348.1; JOINED.
 DR EMBL: D90434; BAA14348.1; JOINED.
 DR EMBL: D90435; BAA14348.1; JOINED.
 DR EMBL: D90436; BAA14348.1; JOINED.
 DR EMBL: D90437; BAA14348.1; JOINED.
 DR EMBL: D90438; BAA14348.1; JOINED.
 DR EMBL: D90439; BAA14348.1; JOINED.
 DR EMBL: D90440; BAA14348.1; JOINED.
 DR EMBL: D90441; BAA14348.1; JOINED.
 DR EMBL: D90442; BAA14348.1; JOINED.
 DR EMBL: D90443; BAA14348.1; JOINED.
 DR EMBL: D90444; BAA14348.1; JOINED.
 DR EMBL: D90445; BAA14348.1; JOINED.
 DR EMBL: D90446; BAA14348.1; JOINED.
 DR EMBL: D90447; BAA14348.1; JOINED.
 DR EMBL: D90448; BAA14348.1; JOINED.
 DR EMBL: D90449; BAA14348.1; JOINED.
 DR EMBL: D90450; BAA14348.1; JOINED.
 DR EMBL: D90451; BAA14348.1; JOINED.
 DR EMBL: D90452; BAA14348.1; JOINED.
 DR EMBL: D90453; BAA14348.1; JOINED.
 DR EMBL: D90454; BAA14348.1; JOINED.
 DR EMBL: D90455; BAA14348.1; JOINED.
 DR EMBL: D90456; BAA14348.1; JOINED.
 DR EMBL: D90457; BAA14348.1; JOINED.
 DR EMBL: D90458; BAA14348.1; JOINED.
 DR EMBL: D90459; BAA14348.1; JOINED.
 DR EMBL: D90460; BAA14348.1; JOINED.
 DR EMBL: D90461; BAA14348.1; JOINED.
 DR EMBL: D90462; BAA14348.1; JOINED.
 DR EMBL: D90463; BAA14348.1; JOINED.
 DR EMBL: D90464; BAA14348.1; JOINED.
 DR EMBL: D90465; BAA14348.1; JOINED.
 DR EMBL: D90466; BAA14348.1; JOINED.
 DR EMBL: D90467; BAA14348.1; JOINED.
 DR EMBL: D90468; BAA14348.1; JOINED.
 DR EMBL: D90469; BAA14348.1; JOINED.
 DR EMBL: D90470; BAA14348.1; JOINED.
 DR EMBL: D90471; BAA14348.1; JOINED.
 DR EMBL: D90472; BAA14348.1; JOINED.
 DR EMBL: D90473; BAA14348.1; JOINED.
 DR EMBL: D90474; BAA14348.1; JOINED.
 DR EMBL: D90475; BAA14348.1; JOINED.
 DR EMBL: D90476; BAA14348.1; JOINED.
 DR EMBL: D90477; BAA14348.1; JOINED.
 DR EMBL: D90478; BAA14348.1; JOINED.
 DR EMBL: D90479; BAA14348.1; JOINED.
 DR EMBL: D90480; BAA14348.1; JOINED.
 DR EMBL: D90481; BAA14348.1; JOINED.
 DR EMBL: D90482; BAA14348.1; JOINED.
 DR EMBL: D90483; BAA14348.1; JOINED.
 DR EMBL: D90484; BAA14348.1; JOINED.
 DR EMBL: D90485; BAA14348.1; JOINED.
 DR EMBL: D90486; BAA14348.1; JOINED.
 DR EMBL: D90487; BAA14348.1; JOINED.
 DR EMBL: D90488; BAA14348.1; JOINED.
 DR EMBL: D90489; BAA14348.1; JOINED.
 DR EMBL: D90490; BAA14348.1; JOINED.
 DR EMBL: D90491; BAA14348.1; JOINED.
 DR EMBL: D90492; BAA14348.1; JOINED.
 DR EMBL: D90493; BAA14348.1; JOINED.
 DR EMBL: D90494; BAA14348.1; JOINED.
 DR EMBL: D90495; BAA14348.1; JOINED.
 DR EMBL: D90496; BAA14348.1; JOINED.
 DR EMBL: D90497; BAA14348.1; JOINED.
 DR EMBL: D90498; BAA14348.1; JOINED.
 DR EMBL: D90499; BAA14348.1; JOINED.
 DR EMBL: D90500; BAA14348.1; JOINED.
 DR EMBL: D90501; BAA14348.1; JOINED.
 DR EMBL: D90502; BAA14348.1; JOINED.
 DR EMBL: D90503; BAA14348.1; JOINED.
 DR EMBL: D90504; BAA14348.1; JOINED.
 DR EMBL: D90505; BAA14348.1; JOINED.
 DR EMBL: D90506; BAA14348.1; JOINED.
 DR EMBL: D90507; BAA14348.1; JOINED.
 DR EMBL: D90508; BAA14348.1; JOINED.
 DR EMBL: D90509; BAA14348.1; JOINED.
 DR EMBL: D90510; BAA14348.1; JOINED.
 DR EMBL: D90511; BAA14348.1; JOINED.
 DR EMBL: D90512; BAA14348.1; JOINED.
 DR EMBL: D90513; BAA14348.1; JOINED.
 DR EMBL: D90514; BAA14348.1; JOINED.
 DR EMBL: D90515; BAA14348.1; JOINED.
 DR EMBL: D90516; BAA14348.1; JOINED.
 DR EMBL: D90517; BAA14348.1; JOINED.
 DR EMBL: D90518; BAA14348.1; JOINED.
 DR EMBL: D90519; BAA14348.1; JOINED.
 DR EMBL: D90520; BAA14348.1; JOINED.
 DR EMBL: D90521; BAA14348.1; JOINED.
 DR EMBL: D90522; BAA14348.1; JOINED.
 DR EMBL: D90523; BAA14348.1; JOINED.
 DR EMBL: D90524; BAA14348.1; JOINED.
 DR EMBL: D90525; BAA14348.1; JOINED.
 DR EMBL: D90526; BAA14348.1; JOINED.
 DR EMBL: D90527; BAA14348.1; JOINED.
 DR EMBL: D90528; BAA14348.1; JOINED.
 DR EMBL: D90529; BAA14348.1; JOINED.
 DR EMBL: D90530; BAA14348.1; JOINED.
 DR EMBL: D90531; BAA14348.1; JOINED.
 DR EMBL: D90532; BAA14348.1; JOINED.
 DR EMBL: D90533; BAA14348.1; JOINED.
 DR EMBL: D90534; BAA14348.1; JOINED.
 DR EMBL: D90535; BAA14348.1; JOINED.
 DR EMBL: D90536; BAA14348.1; JOINED.
 DR EMBL: D90537; BAA14348.1; JOINED.
 DR EMBL: D90538; BAA14348.1; JOINED.
 DR EMBL: D90539; BAA14348.1; JOINED.
 DR EMBL: D90540; BAA14348.1; JOINED.
 DR EMBL: D90541; BAA14348.1; JOINED.
 DR EMBL: D90542; BAA14348.1; JOINED.
 DR EMBL: D90543; BAA14348.1; JOINED.
 DR EMBL: D90544; BAA14348.1; JOINED.
 DR EMBL: D90545; BAA14348.1; JOINED.
 DR EMBL: D90546; BAA14348.1; JOINED.
 DR EMBL: D90547; BAA14348.1; JOINED.
 DR EMBL: D90548; BAA14348.1; JOINED.
 DR EMBL: D90549; BAA14348.1; JOINED.
 DR EMBL: D90550; BAA14348.1; JOINED.
 DR EMBL: D90551; BAA14348.1; JOINED.
 DR EMBL: D90552; BAA14348.1; JOINED.
 DR EMBL: D90553; BAA14348.1; JOINED.
 DR EMBL: D90554; BAA14348.1; JOINED.
 DR EMBL: D90555; BAA14348.1; JOINED.
 DR EMBL: D90556; BAA14348.1; JOINED.
 DR EMBL: D90557; BAA14348.1; JOINED.
 DR EMBL: D90558; BAA14348.1; JOINED.
 DR EMBL: D90559; BAA14348.1; JOINED.
 DR EMBL: D90560; BAA14348.1; JOINED.
 DR EMBL: D90561; BAA14348.1; JOINED.
 DR EMBL: D90562; BAA14348.1; JOINED.
 DR EMBL: D90563; BAA14348.1; JOINED.
 DR EMBL: D90564; BAA14348.1; JOINED.
 DR EMBL: D90565; BAA14348.1; JOINED.
 DR EMBL: D90566; BAA14348.1; JOINED.
 DR EMBL: D90567; BAA14348.1; JOINED.
 DR EMBL: D90568; BAA14348.1; JOINED.
 DR EMBL: D90569; BAA14348.1; JOINED.
 DR EMBL: D90570; BAA14348.1; JOINED.
 DR EMBL: D90571; BAA14348.1; JOINED.
 DR EMBL: D90572; BAA14348.1; JOINED.
 DR EMBL: D90573; BAA14348.1; JOINED.
 DR EMBL: D90574; BAA14348.1; JOINED.
 DR EMBL: D90575; BAA14348.1; JOINED.
 DR EMBL: D90576; BAA14348.1; JOINED.
 DR EMBL: D90577; BAA14348.1; JOINED.
 DR EMBL: D90578; BAA14348.1; JOINED.
 DR EMBL: D90579; BAA14348.1; JOINED.
 DR EMBL: D90580; BAA14348.1; JOINED.
 DR EMBL: D90581; BAA14348.1; JOINED.
 DR EMBL: D90582; BAA14348.1; JOINED.
 DR EMBL: D90583; BAA14348.1; JOINED.
 DR EMBL: D90584; BAA14348.1; JOINED.
 DR EMBL: D90585; BAA14348.1; JOINED.
 DR EMBL: D90586; BAA14348.1; JOINED.
 DR EMBL: D90587; BAA14348.1; JOINED.
 DR EMBL: D90588; BAA14348.1; JOINED.
 DR EMBL: D90589; BAA14348.1; JOINED.
 DR EMBL: D90590; BAA14348.1; JOINED.
 DR EMBL: D90591; BAA14348.1; JOINED.
 DR EMBL: D90592; BAA14348.1; JOINED.
 DR EMBL: D90593; BAA14348.1; JOINED.
 DR EMBL: D90594; BAA14348.1; JOINED.
 DR EMBL: D90595; BAA14348.1; JOINED.
 DR EMBL: D90596; BAA14348.1; JOINED.
 DR EMBL: D90597; BAA14348.1; JOINED.
 DR EMBL: D90598; BAA14348.1; JOINED.
 DR EMBL: D90599; BAA14348.1; JOINED.
 DR EMBL: D90600; BAA14348.1; JOINED.
 DR EMBL: D90601; BAA14348.1; JOINED.
 DR EMBL: D90602; BAA14348.1; JOINED.
 DR EMBL: D90603; BAA14348.1; JOINED.
 DR EMBL: D90604; BAA14348.1; JOINED.
 DR EMBL: D90605; BAA14348.1; JOINED.
 DR EMBL: D90606; BAA14348.1; JOINED.
 DR EMBL: D90607; BAA14348.1; JOINED.
 DR EMBL: D90608; BAA14348.1; JOINED.
 DR EMBL: D90609; BAA14348.1; JOINED.
 DR EMBL: D90610; BAA14348.1; JOINED.
 DR EMBL: D90611; BAA14348.1; JOINED.
 DR EMBL: D90612; BAA14348.1; JOINED.
 DR EMBL: D90613; BAA14348.1; JOINED.
 DR EMBL: D90614; BAA14348.1; JOINED.
 DR EMBL: D90615; BAA14348.1; JOINED.
 DR EMBL: D90616; BAA14348.1; JOINED.
 DR EMBL: D90617; BAA14348.1; JOINED.
 DR EMBL: D90618; BAA14348.1; JOINED.
 DR EMBL: D90619; BAA14348.1; JOINED.
 DR EMBL: D90620; BAA14348.1; JOINED.
 DR EMBL: D90621; BAA14348.1; JOINED.
 DR EMBL: D90622; BAA14348.1; JOINED.
 DR EMBL: D90623; BAA14348.1; JOINED.
 DR EMBL: D90624; BAA14348.1; JOINED.
 DR EMBL: D90625; BAA14348.1; JOINED.
 DR EMBL: D90626; BAA14348.1; JOINED.
 DR EMBL: D90627; BAA14348.1; JOINED.
 DR EMBL: D90628; BAA14348.1; JOINED.
 DR EMBL: D90629; BAA14348.1; JOINED.
 DR EMBL: D90630; BAA14348.1; JOINED.
 DR EMBL: D90631; BAA14348.1; JOINED.
 DR EMBL: D90632; BAA14348.1; JOINED.
 DR EMBL: D90633; BAA14348.1; JOINED.
 DR EMBL: D90634; BAA14348.1; JOINED.
 DR EMBL: D90635; BAA14348.1; JOINED.
 DR EMBL: D90636; BAA14348.1; JOINED.
 DR EMBL: D90637; BAA14348.1; JOINED.
 DR EMBL: D90638; BAA14348.1; JOINED.
 DR EMBL: D90639; BAA14348.1; JOINED.
 DR EMBL: D90640; BAA14348.1; JOINED.
 DR EMBL: D90641; BAA14348.1; JOINED.
 DR EMBL: D90642; BAA14348.1; JOINED.
 DR EMBL: D90643; BAA14348.1; JOINED.
 DR EMBL: D90644; BAA14348.1; JOINED.
 DR EMBL: D90645; BAA14348.1; JOINED.
 DR EMBL: D90646; BAA14348.1; JOINED.
 DR EMBL: D90647; BAA14348.1; JOINED.
 DR EMBL: D90648; BAA14348.1; JOINED.
 DR EMBL: D90649; BAA14348.1; JOINED.
 DR EMBL: D90650; BAA14348.1; JOINED.
 DR EMBL: D90651; BAA14348.1; JOINED.
 DR EMBL: D90652; BAA14348.1; JOINED.
 DR EMBL: D90653; BAA14348.1; JOINED.
 DR EMBL: D90654; BAA14348.1; JOINED.
 DR EMBL: D90655; BAA14348.1; JOINED.
 DR EMBL: D90656; BAA14348.1; JOINED.
 DR EMBL: D90657; BAA14348.1; JOINED.
 DR EMBL: D90658; BAA14348.1; JOINED.
 DR EMBL: D90659; BAA14348.1; JOINED.
 DR EMBL: D90660; BAA14348.1; JOINED.
 DR EMBL: D90661; BAA14348.1; JOINED.
 DR EMBL: D9066

FT CONFLICT 479 479 V -> L (IN REF. 2).
FT CONFLICT 564 564 R -> H (IN REF. 3).
SQ SEQUENCE 728 AA; 82944 MW; A0381FC497534328 CRC64;

Query Match 70.0%; Score 35; DB 1; Length 728;
Best Local Similarity 50.0%; Pred. No. 24;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VVRIGHLYIL 10
Db 631 LRRVAHLXIM 640

RESULT 8
HGF_RAT STANDARD; PRT; 728 AA.
ID HGF_RAT
AC P17945;
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HEPATOCYTE GROWTH FACTOR PRECURSOR (SCATTER FACTOR) (SF)
DE (HEPATOPOEITIN A).
GN HGF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RM [1]
RM SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC STRAIN-WISTAR; TISSUE=Liver;
RX MEDLINE=90222197; PubMed=2139229;
RA Toshio K., Hagiya M., Nishizawa T., Seki T., Shimomishi M.,
RA Shimizu S., Nakamura T.;
RT "Deduced primary structure of rat hepatocyte growth factor and
RT expression of the mRNA in rat tissues.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3200-3204(1990).
RM [2]
RM SEQUENCE FROM N.A.
RC STRAIN-WISTAR; TISSUE=Liver;
RX MEDLINE=91031482; PubMed=2146117;
RA Okajima A., Miyazawa K., Kitamura N.;
RT "Primary structure of rat hepatocyte growth factor and induction of
RT its mRNA during liver regeneration following hepatic injury.";
RL Eur. J. Biochem. 193:375-381(1990).
CC -1- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
CC HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPIC FACTOR, AND ACTS
CC AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.
CC -1- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
CC DISULFIDE BOND.
CC -1- SIMILARITY: CONTAINS 4 KRINGLE REGIONS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY. PLASMINOGEN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D90102; BAA14133.1; -;
DR EMBL: X54400; CAA38266.1; -;
DR PIR: S13211; S13211.
DR PIR: A35644; A35644.
DR HSSP: P14210; 2HGF.
DR MEROPS: S01.978; -;
DR InterPro: IPR000001; -;
DR InterPro: IPR001254; -;
DR InterPro: IPR001314; -;
DR InterPro: IPR003014; -;

DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00051; kringle; 4.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00018; KRINGLE.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PROSITE: PS00021; KRINGLE_1; 4.
DR PROSITE: PS0070; KRINGLE_2; 4.
KW Growth factor; Kringle; Glycoprotein; Serine protease homolog;
KW Signal.
FT SIGNAL 1 32 BY SIMILARITY.
FT CHAIN 33 485 HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.
FT CHAIN 496 728 HEPATOCYTE GROWTH FACTOR BETA CHAIN.
FT MOD_RES 33 33 PYRROLIDONE CARBOXYLIC ACID
(BY SIMILARITY).
FT DOMAIN 33 128 PAP.
FT DOMAIN 129 207 KRINGLE 1.
FT DOMAIN 212 289 KRINGLE 2.
FT DOMAIN 306 384 KRINGLE 3.
FT DOMAIN 392 470 KRINGLE 4.
FT DOMAIN 496 728 SERINE PROTEASE-LIKE.
FT DISULFID 71 97 BY SIMILARITY.
FT DISULFID 75 85 INTERCHAIN (BY SIMILARITY).
FT DISULFID 488 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 656 656 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 728 AA; 82905 MW; 3E0BF1F6ADCEDF CRC64;

Query Match 70.0%; Score 35; DB 1; Length 728;
Best Local Similarity 50.0%; Pred. No. 24;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VVRIGHLYIL 10
Db 631 LRRVAHLXIM 640

RESULT 9
AT1D_HUMAN STANDARD; PRT; 933 AA.
ID AT1D_HUMAN
AC P98198;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE POTENTIAL PHOSPHOLIPID-TRANSPORTING ATPASE ID (EC 3.6.1.-) (FRAGMENT).
GN KIAA1137.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RM [1]
RM SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039618; PubMed=10574461;
RA Hirosewa M., Nagase T., Ishikawa K.-I., Kikuno R., Nomura N.,
RA Ohara O.;
RT "Characterization of cDNA clones selected by the Genemark analysis
RT from size-fractionated cDNA libraries from human brain.";
RL DNA Res. 6:329-336(1999).
RM [2]
RM SEQUENCE OF 836-933 FROM N.A.
RC TISSUE=Testis;
RA Othenaelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + ORTHOPHOSPHATE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2
CC ATPASES). SUBFAMILY IV.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB032963; BAA6451.1; -
 DR EMBL: AL137537; CAB70799.1; -
 DR PROSITE: PS00154; ATPASE_E1_E2; 1.
 KW Hydrolyase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;
 KW Multigene family.
 FT DOMAIN 1 19 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 20 41 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 42 70 POTENTIAL.
 FT TRANSMEM 71 92 POTENTIAL.
 FT DOMAIN 93 613 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 614 634 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 635 646 POTENTIAL.
 FT TRANSMEM 647 666 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 667 696 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 697 718 POTENTIAL.
 FT DOMAIN 719 732 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 733 755 POTENTIAL.
 FT DOMAIN 756 761 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 762 782 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 803 827 POTENTIAL.
 FT TRANSMEM 828 933 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 933 135 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 157 557 MAGNESIUM (BY SIMILARITY).
 FT METAL 561 561 MAGNESIUM (BY SIMILARITY).
 SQ SEQUENCE 933 AA; 105849 MW; CAF60E3574275AC1 CRC64;
 OY 1 VVRIGHLYIT 10
 Db 91 VVRIGHLYIT 100
 Query Match 70.0%; Score 35; DB 1; Length 933;
 Best Local Similarity 50.0%; Pred. No. 31;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

RESULT 10
 ID CDSL_SOLTU STANDARD; PRT; 424 AA.
 AC 004940;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PHOSPHATIDATE CYTIDYLTRANSFERASE (EC 2.7.7.41) (CDP-DIGLYCERIDE
 DE SYNTHETASE) (CDP-DIGLYCERIDE PYROPHOSPHORYLASE) (CDP-DIACYLGlycerol
 DE SYNTHETASE) (CDP-CTP:PHOSPHATIDATE CYTIDYLTRANSFERASE) (CDP-DAG
 DE SYNTHETASE) (CDP-DG SYNTHETASE).
 GN CDSL.
 OS Solanum tuberosum (Potato)
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; eunasterids I;
 OC Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4113;
 RX MEDLINE=97239925; PubMed=9085581;
 RA Kopka J., Ludwig M., Mueller-Roeber B.;
 RT "Complementary DNAs encoding eukaryotic-type cytidine-5'-diphosphate-
 RT diacylglycerol synthases of two plant species.";
 RL Plant Physiol. 113:997-1002(1997).
 CC -1- FUNCTION: MAY BE INVOLVED IN THE SYNTHESIS OF MINOR PHOSPHOLIPIDS
 CC AND IN MODULATION OF IP3-MEDIATED SIGNAL TRANSDUCTION.
 CC -1- CATALYTIC ACTIVITY: CTP + PHOSPHATIDATE = PYROPHOSPHATE +

CC CDP-DIACYLGLYCEROL.
 CC -1- PATHWAY: PHOSPHOLIPID BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: ROOTS AND SINK LEAVES.
 CC -1- SIMILARITY: BELONGS TO THE CDS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X91909; CAA63004.1; -
 DR InterPro: IPR000374; -
 DR Pfam: PF01148; Cytidylyltrans; 1.
 DR PROSITE: PS01315; CDS; 1.
 KW Transferase; Nucleotidyltransferase; Phospholipid biosynthesis;
 KW Transmembrane.
 FT TRANSMEM 60 80 POTENTIAL.
 FT TRANSMEM 149 169 POTENTIAL.
 FT TRANSMEM 206 226 POTENTIAL.
 FT TRANSMEM 246 266 POTENTIAL.
 FT TRANSMEM 321 341 POTENTIAL.
 FT DOMAIN 341 424 POLY-ARG.
 SQ SEQUENCE 424 AA; 49178 MW; 43B556B12F082638 CRC64;
 OY 1 VVRIGHLYIT 9
 Db 66 IIVGHLYIT 74
 Query Match 68.0%; Score 34; DB 1; Length 424;
 Best Local Similarity 55.6%; Pred. No. 22;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

RESULT 11
 ID MAG3_HUMAN STANDARD; PRT; 314 AA.
 AC P43357;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE MELANOMA-ASSOCIATED ANTIGEN 3 (MAGE-3 ANTIGEN) (ANTIGEN M22-D).
 GN MAGE3 OR MAGE3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND MUTAGENESIS.
 RC TISSUE-Blood;
 RX MEDLINE=94157413; PubMed=8113684;
 RA Gaugler B., van den Eynde B., van der Bruggen P., Romero P.,
 RA Gattorfo J.J., de Plaen E., Lethe B., Brasseur F., Boon T.;
 RT "Human gene MAGE-3 codes for an antigen recognized on a melanoma by
 RT autologous cytolytic T lymphocytes.";
 RL J. Exp. Med. 179:921-930(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Skin;
 RX MEDLINE=94311935; PubMed=8037761;
 RA Ding M., Beck R.J., Keller C.J., Penton R.G.;
 RT "Cloning and analysis of MAGE-1-related genes.";
 RL Biochem. Biophys. Res. Commun. 202:549-553(1994).
 CC -1- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
 CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
 CC PROGRESSION. ANTIGEN RECOGNIZED ON A MELANOMA BY AUTOLOGOUS
 CC CYTOLYTIC T LYMPHOCYTES.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
 CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG

```

CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
CC FOR TESTES AND PLACENTA. NEVER EXPRESSED IN KIDNEY TUMORS,
CC LEUKEMIAS AND LYMPHOMAS.
CC -1- SIMILARITY: BELONGS TO THE MAGE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U03735; AAA17446.1; -.
DR MIM; 300174; -.
DR InterPro; IPR002190; -.
DR Pfam; PF01454; MAGE; 1.
KW Antigen; Multigene family; Tumor antigen.
FT DOMAIN 40 43 POLY-SER.
FT MUTAGEN 170 170 D->A: ABOLISHES HLA-A1 BINDING.
FT MUTAGEN 176 176 Y->A: ABOLISHES HLA-A1 BINDING.
SQ SEQUENCE 314 AA; 34747 MW; 3F5E13D1C946A1 CRC64;

Query Match
Best Local Similarity 100.0%; Score 33; DB 1; Length 314;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 IGHLY1 9
Db 172 IGHLY1 177

RESULT 12
TNSE_ECOLI STANDARD; PRT; 538 AA.
AC P05845;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE TRANSPOSON TNY7 TRANSPOSITION PROTEIN TNSE (PROTEIN D).
GN TNSE.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90192166; PubMed=2156235;
RA Flores C., Gadi M.I., Lichtenstein C.;
RT "DNA sequence analysis of five genes; tnsA, B, C, D and E, required
RT for Tny7 transposition.";
RL Nucleic Acids Res. 18:901-911(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87040763; PubMed=3022239;
RA Smith G.M., Jones P.;
RT "Tny7 transposition: a multigene process. Identification of a
RT regulatory gene product.";
RL Nucleic Acids Res. 14:7915-7927(1986).
CC -1- FUNCTION: TNSABC + TNSD PROMOTE HIGH-FREQUENCY INSERTION OF TNY7
CC INTO A SPECIFIC TARGET SITE KNOWN AS ATT-TNY7 WHEREAS TNSABC + TNSE
CC PROMOTE LOW-FREQUENCY INSERTION INTO MANY DIFFERENT SITES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17693; CAA35687.1; -.

```

```

DR EMBL; X04534; CAB56509.1; -.
DR PIR; A25543; OOECD7.
DR PIR; S12641; S12641.
DR PIR; S06770; S06770.
KW Transposable element; Transposition; DNA recombination; DNA-binding.
FT DNA_BIND 311 330 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 538 AA; 61211 MW; DE34A3F1A1A185B CRC64;

Query Match
Best Local Similarity 75.0%; Score 33; DB 1; Length 538;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VVRIGLY 8
Db 13 VVRIGLY 20

RESULT 13
ETFD_PIG STANDARD; PRT; 607 AA.
AC P59331;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE ELECTRON TRANSFER FLAVOPROTEIN-UBIQUINONE OXIDOREDUCTASE PRECURSOR
DE (EC 1.5.5.1) (ETFD-00) (ETFD-UBIQUINONE OXIDOREDUCTASE) (ETFD
DE DEHYDROGENASE) (ELECTRON-TRANSFERRING-FLAVOPROTEIN DEHYDROGENASE)
DE (FRAGMENT).
GN ETFD.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Fetal liver.
RX MEDLINE=94139702; PubMed=8306995;
RA Goodman S.I., Axtell K.M., Bindoff L.A., Beard S.E., Gill R.E.,
RA Freeman F.E.;
RT "Molecular cloning and expression of a cDNA encoding human electron
RT transfer flavoprotein-ubiquinone oxidoreductase.";
RL Eur. J. Biochem. 219:277-286(1994).
CC -1- FUNCTION: ACCEPTS ELECTRONS FROM ETF AND REDUCES UBIQUINONE.
CC -1- CATALYTIC ACTIVITY: REDUCED ETF + UBIQUINONE = ETF + UBIQUINOL.
CC -1- COFACTOR: FAD AND A 4FE-4S CLUSTER.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE (BY
CC SIMILARITY).
KW Oxidoreductase; Electron transport; Flavoprotein; FAD; Iron-sulfur;
KW 4Fe-4S; Mitochondrion; Transit peptide; Ubiquinone; Transmembrane.
FT NOD_TER 1 1
FT TRANSIT 1 23
FT CHAIN 24 607
FT NP_BIND 61 75
FT TRANSMEM 448 468
FT METAL 551 551 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 576 576 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 579 579 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 582 582 IRON-SULFUR (4FE-4S) (POTENTIAL).
SQ SEQUENCE 607 AA; 67590 MW; 2435780774C6FDE CRC64;

Query Match
Best Local Similarity 85.7%; Score 33; DB 1; Length 607;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVRIGLY 7
Db 163 VVRIGLY 169

```

```

RESULT 14
GCKR_XENLA STANDARD; PRT; 619 AA.
ID GCKR_XENLA
AC Q91754;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GLUCOKINASE REGULATORY PROTEIN (GLUCOKINASE REGULATOR).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8335;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95010134; PubMed=7925465;
RA Velga-Da-Cunha M., Delneux M., Matelet N., van Schaftingen E.;
RT "Cloning and expression of a Xenopus liver cDNA encoding a fructose-
RL phosphate-insensitive regulatory protein of glucokinase."
CC Bur. J. Biochem. 225:43-51(1994).
CC -1- FUNCTION: INHIBITS GLUCOKINASE BY FORMING AN INACTIVE COMPLEX WITH
CC THIS ENZYME.
CC -1- SIMILARITY: BELONGS TO THE SIS FAMILY. GCKR SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X80901; CAA56863.1; -
DR InterPro: IPR001347; -
DR InterPro: IPR001741; -
DR Pfam: PF01380; SIS; 1.
DR PROSITE: PS01272; GCKR; 1.
SQ SEQUENCE 619 AA; 68738 MM; 41B72C1981D1BA52 CRC64;

Query Match 66.0%; Score 33; DB 1; Length 619;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GHLXYIL 10
|||||
DB 327 GHLXYIL 332

RESULT 15
SOMA_CRONO STANDARD; PRT; 190 AA.
ID SOMA_CRONO
AC P55755;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SOMATOTROPIN (GROWTH HORMONE).
OS Crocodylus novaeguineae (Crocodylle).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Crocodylinae; Crocodylus.
OX NCBI_TaxID=8503;
RN [1]
RP SEQUENCE.
RC TISSUE=pituitary;
RX MEDLINE=95354920; PubMed=7628683;
RA Noso T., Lance V.A., Kawauchi H.;
RT "Complete amino acid sequence of crocodile growth hormone.";
RL Gen. Comp. Endocrinol. 98:244-252(1995).
CC -1- FUNCTION: GROWTH HORMONE PLAYS AN IMPORTANT ROLE IN GROWTH
CC CONTROL AND INVOLVED IN THE REGULATION OF SEVERAL ANABOLIC
CC PROCESSES.
CC -1- SUBCELLULAR LOCATION: SECRETED.

```

```

CC -1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
DR HSSP: P01246; IBST.
DR InterPro: IPR001400; -
DR Pfam: PF00103; hormone; 1.
DR PRINTS: PR00836; SOMATOTROPIN.
DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
KM Hormone; Pituitary.
FT DISULFID 52 163 BY SIMILARITY.
FT DISULFID 180 188 BY SIMILARITY.
SQ SEQUENCE 190 AA; 22008 MM; 7D8EC7940A50E35A CRC64;

Query Match 64.0%; Score 32; DB 1; Length 190;
Best Local Similarity 60.0%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVRIGHLTYIL 10
|:| |||:|
DB 14 VLRQHLTYLL 23

```

Search completed: June 20, 2001, 14:09:31
Job time: 328 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2001, 14:10:22 ; Search time 47.6 Seconds

(without alignments)
27.795 Million cell updates/sec

Title: US-09-692-401-5

Sequence: 1 VVRIGHLYTL 10

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_PROTOZOA:*
12: SP_UNCLASSIFIED:*
13: SP_VERTEBRATE:*
14: SP_VIRUS:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	100.0	314	4 Q9NSD3	Q9NSD3 homo sapien
2	38	76.0	233	10 Q9S181	Q9S181 arabidopsis
3	36	72.0	423	10 Q49639	Q49639 arabidopsis
4	36	72.0	1310	5 Q9VVF8	Q9VVF8 drosophila
5	35	70.0	234	2 Q9S1R5	Q9S1R5 streptomyces
6	35	70.0	305	2 Q45454	Q45454 bacillus su
7	35	70.0	325	2 Q34558	Q34558 bacillus su
8	35	70.0	325	9 Q64101	Q64101 bacterioph
9	35	70.0	520	10 Q65815	Q65815 helianthus
10	35	70.0	531	2 Q9RMW0	Q9RMW0 bacillus an
11	35	70.0	637	2 Q06769	Q06769 mycobacteri
12	35	70.0	723	4 Q9UDU6	Q9UDU6 homo sapien
13	35	70.0	808	10 Q9LNV7	Q9LNV7 arabidopsis
14	35	70.0	1863	11 Q9JLQ1	Q9JLQ1 mus musculu
15	34	68.0	139	2 Q49122	Q49122 methylobact
16	34	68.0	304	2 Q9WYQ3	Q9WYQ3 thermotoga
17	34	68.0	340	4 Q9GZV7	Q9GZV7 homo sapien
18	34	68.0	341	11 Q9ESM3	Q9ESM3 mus musculu
19	34	68.0	341	11 Q9ESM2	Q9ESM2 rattus norv

20	34	68.0	363	2 Q9EVH2	Q9EVH2 buchnera ap
21	34	68.0	438	10 Q9S217	Q9S217 arabidopsis
22	34	68.0	1235	13 Q9PTG7	Q9PTG7 xenopus lae
23	34	68.0	1244	10 Q82607	Q82607 arabidopsis
24	34	68.0	1288	10 Q9STV0	Q9STV0 arabidopsis
25	34	68.0	1461	10 Q9S1M3	Q9S1M3 arabidopsis
26	33	66.0	101	2 Q9PHS1	Q9PHS1 xylella fas
27	33	66.0	168	2 Q9PD80	Q9PD80 xylella fas
28	33	66.0	301	2 Q52770	Q52770 pseudomonas
29	33	66.0	313	1 Q28287	Q28287 archaeoglob
30	33	66.0	324	2 Q9Z4K0	Q9Z4K0 streptomyces
31	33	66.0	337	1 Q9YFK2	Q9YFK2 aeropyrum p
32	33	66.0	338	1 Q9UWX3	Q9UWX3 sulfolobus
33	33	66.0	361	1 Q9UWV2	Q9UWV2 sulfolobus
34	33	66.0	444	2 Q85748	Q85748 treponema d
35	33	66.0	604	5 Q9V3D9	Q9V3D9 drosophila
36	33	66.0	663	2 Q9RWV5	Q9RWV5 delnoccocus
37	33	66.0	1112	1 Q9UVR7	Q9UVR7 pyrococcus
38	33	66.0	1232	10 Q23998	Q23998 hordeum vul
39	32	64.0	110	2 Q49296	Q49296 mycoplasma
40	32	64.0	154	6 Q9M2A9	Q9M2A9 ovis aries
41	32	64.0	176	2 Q87064	Q87064 vibrio chol
42	32	64.0	191	10 Q9ZVX9	Q9ZVX9 arabidopsis
43	32	64.0	196	10 Q9LP00	Q9LP00 arabidopsis
44	32	64.0	199	2 Q88119	Q88119 listeria mo
45	32	64.0	204	13 Q9PRH7	Q9PRH7 oreochromis

ALIGNMENTS

RESULT 1

ID Q9NSD3 PRELIMINARY: PRT: 314 AA.

AC Q9NSD3: 01-OCT-2000 (TREMBL) 15, Created

DT 01-OCT-2000 (TREMBL) 15, Last sequence update

DE 01-MAR-2001 (TREMBL) 16, Last annotation update

DE MEINOMA ANTIGEN FAMILY A12.

GN MAGFAL2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Mallon A.M., Platzer M., Bates R., Gloeckner G., Botcherby M.,

RA Nordstiek G., Strivens M.A., Kioschis P., Dangel A., Cunningham D.,

RA Straw R., Weston P., Hunter C., Gilbert M., Fernando S., Goodall K.,

RA Kerry G., Greystrom J.S., Clark D., Goeldes M., Blechschmidt K.,

RA Rhodes M., Denny P., Mundy C.R., Miller W., Poustka A., Herman G.E.,

RT "Comparative genome sequence analysis of the Bpa/Stc region in mouse

RT and man.,"

RL Genome Res. 0:0-0(2000).

DR EMBL: U82671; AAF4789.1; -.

DR InterPro: IPR002190; -.

DR Pfam: PF01434; MAGF; 1.

SQ SEQUENCE 314 AA: 34836 MW: 7E0DF7CECD8F6568 CRC64;

Query Match 100.0%; Score 50; DB 4; Length 314;

Best local similarity 100.0%; Pred. No. 0.097;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVRIGHLYTL 10

DB 169 VVRIGHLYTL 178

RESULT 2
ID Q9S181 PRELIMINARY: PRT: 233 AA.

AC Q9S181;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE F33N19.5.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Walker M., Shinn P., Brooks S., Buehler E., Chao Q., Dunn P., Khan S.,
 RA Kim C., Altieri H., Araujo R., Conn L., Conway A.B., Gonzalez A.,
 RA Hansen N.F., Hultzer L., Kremenetskaia I., Lenz C., Li J., Liu S.,
 RA Luros S., Rowley D., Schwartz J., Tortum M., Vysotskaia V., Yu G.,
 RA Davis R.W., Federspiel N.A., Theologis A., Ecker J.R.,
 RT "Genomic sequence for Arabidopsis thaliana BAC F23N19 from chromosome
 1.";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC007190; AAF19536.1; -
 SO SEQUENCE 233 AA; 26150 MW; 7D292F68D261C6EE CRC64;

Query Match 76.0%; Score 38; DB 10; Length 233;
 Best Local Similarity 75.0%; Pred. No. 15;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVRIGHLY 8
 Db 176 IVRIGHLY 183

RESULT 3
 ID 049639 PRELIMINARY; PRT; 423 AA.
 AC 049639;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE CDP-DIACYLGLYCEROL SYNTHETASE-LIKE PROTEIN.
 GN T10114.170 OR ATG422340.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bayan M., Murphy G., Drost L., Hall C., Hudson S., Ridley P.,
 RA Bancroft I., Mewes H.W., Mayer K., Schueller C.,
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL021712; CAAL6784.1; -
 DR EMBL: AL161557; CAB79189.1; -
 DR EMBL: 28363; Atarch;2834;28363.
 DR Mendel; 28363; Atarch;2834;28363.
 DR InterPro: IPR000374; -
 DR Pfam: PF01148; Cytidylyltrans; 1.
 DR PROSITE: PS01315; CDS; 1.
 SO SEQUENCE 423 AA; 48247 MW; 02776A12440FA407 CRC64;

Query Match 72.0%; Score 36; DB 10; Length 423;
 Best Local Similarity 66.7%; Pred. No. 66;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VVRIGHLY 9
 Db 66 IVRIGHLY 74

RESULT 4
 ID Q9VVF8 PRELIMINARY; PRT; 1310 AA.
 AC Q9VVF8;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE CG7692 PROTEIN.
 GN CG7692.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-BERKELEY;
 RC MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Abmayant A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.A., Bouck J., Brokstein P., Brotler P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Daventport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kammel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclio J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sidenkiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195 (2000).
 DR EMBL: AE003524; AAF49353.1; -
 DR FlyBase: FBgn0036714; CG7692.
 SO SEQUENCE 1310 AA; 150503 MW; 2AF07754E8452430 CRC64;

Query Match 72.0%; Score 36; DB 5; Length 1310;
 Best Local Similarity 60.0%; Pred. No. 2.2e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VVRIGHLY 10
 Db 216 ITRIGHLY 225

RESULT 5
 Q9S1R5


```

ID Q9S1R5      PRELIMINARY;      PRT;      234 AA.
AC Q9S1R5;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
DE PUTATIVE MEMBRANE PROTEIN.
GN SCJ9A.05C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Harris D.;
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE-97000351; PubMed-8843436;
RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmid maps and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL: AL109972; CAB53266.1; -
SO SEQUENCE 234 AA; 25619 MW; AAFERBLICDD014319 CRC64;

Query Match      70.0%; Score 35; DB 2; Length 234;
Best Local Similarity 55.6%; Pred. No. 56;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 VRIGHLYL 10
Db 148 LEIGHMYVL 156

RESULT 6
ID Q45454      PRELIMINARY;      PRT;      305 AA.
AC Q45454;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE PROBABLY POSITIONED IN OPERON WITH DOWNSTREAM ORF5.
OS Bacillus subtilis.
OG Plasmid pTA1060.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-IF03022;
RX MEDLINE-95206941; PubMed-7899081;
RA Meijer W.J., Venema G., Bron S.;
RT Characterization of single strand origins of cryptic rolling-circle
RT plasmids from Bacillus subtilis.
RL Nucleic Acids Res. 23:612-619(1995).
DR EMBL: U32380; AAC44422.1; -
KM Plasmid.
SO SEQUENCE 305 AA; 35592 MW; DB410B2A7BC51B22 CRC64;

Query Match      70.0%; Score 35; DB 2; Length 305;
Best Local Similarity 66.7%; Pred. No. 73;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 2 VRIGHLYL 10
Db 197 VILGHMYVL 205

RESULT 7
ID Q34558      PRELIMINARY;      PRT;      325 AA.
AC Q34558;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE YOPR PROTEIN.
GN YOPR.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RX MEDLINE-98044033; PubMed-9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolognin A., Borchert S.,
RA Borst R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,
RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Ertlan K.D., Errington J., Fabre C., Ferrati E., Foulger D.,
RA Fitzer C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grand G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koester P., Konungstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Portwolik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scottone F.,
RA Sekiguchi U., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzengger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: 289114; CAB13971.1; -
DR EMBL: 289115; CAB13997.1; -
SO SEQUENCE 325 AA; 37571 MW; EB73F9970BC5AC3 CRC64;

Query Match      70.0%; Score 35; DB 2; Length 325;
Best Local Similarity 75.0%; Pred. No. 79;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VVRIGHLY 8
Db 256 IVRSGLHY 263

RESULT 8

```

064101 PRELIMINARY; PRT: 325 AA.
 AC 064101;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)
 DE HYPOHETICAL 37.6 KDA PROTEIN.
 GN YOPR.
 OS Bacteriophage SPBc2.
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.
 NCBI_TaxID=6797;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lazarevic V., Duesterhoeft A., Solido B., Hilbert H., Mauel C.,
 RU Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF020713; AAC13061.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 325 AA; 37571 MW; EB73E9970BC5AC63 CRC64;

Query Match 70.0%; Score 35; DB 9; Length 325;
 Best Local Similarity 75.0%; Pred. No. 79;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VVRIGHLY 8
 :|||:|
 DB 256 IVRSGHLY 263

RESULT 9
 ID 065815 PRELIMINARY; PRT: 520 AA.
 AC 065815: 065814;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE CYTOCHROME P450 81B1 (EC 1.14.13.53) (ISOFLAVONE 2'-HYDROXYLASE).
 GN CYP81B1 OR CYP81B1 OR CYP81B1C OR CYP81B1S.
 OS Helianthus tuberosus (Jerusalem artichoke).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons: core eudicots; Asteridae;
 OC eumasterids II; Asterales; Asteraceae; Asteroideae; Helianthaceae;
 OC Helianthus.
 NCBI_TaxID=4233;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=CV. BLANC COMMUN; TISSUE=TUBER;
 RX MEDLINE=98184826; PubMed=9516419;
 RA Cabello-Hurtado F., Batard Y., Salaun J.-P., Durst F., Pinot F.,
 RA Merck-Reichhart D.;
 RT "Cloning, expression in yeast, and functional characterization of
 CYP81B1, a plant cytochrome P450 that catalyzes in-chain hydroxylation
 of fatty acids.";
 RL J. Biol. Chem. 273:7260-7267(1998).
 CC -1- FUNCTION: CATALYZES THE HYDROXYLATION OF ISOFLAVONES, DAIDZEIN AND
 FORMONONETIN, TO YIELD 2'-HYDROXYISOFLAVONES, 2'-HYDROXYDAIDZEIN,
 AND 2'-HYDROXYFORMONONETIN, RESPECTIVELY (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: FORMONONETIN + NADPH + O(2) =
 2-HYDROXYFORMONONETIN + NADP(+) + H(2)O.
 CC -1- COFACTOR: THE ENZYME IS NADPH-DEPENDENT.
 CC -1- ENZYME REGULATION: ENZYME ACTIVITY IS INDUCED BY MNC12 OR
 AMINOPURINE.
 CC -1- PATHWAY: INVOLVED IN THE BIOSYNTHESIS OF ISOFLAVONOID-DERIVED
 ANTICANCERIAL PRODUCTS. TWO FORMS: CYP81B1 (SHOWN HERE) AND
 CC -1- CYP81B1S, ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL: AJ000478; CA04117.1; -
 DR EMBL: AJ000477; CA04116.1; -
 DR InterPro: IPR001128; -
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 RC STRAIN=H37RV;

DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN.1.
 KW Oxidoreductase; Monooxygenase; Membrane; Heme; Alternative splicing;
 NCBI_TaxID=4233;
 GN NADP.
 FT BINDING 456 456 HEME (BY SIMILARITY).
 FT VARSPLIC 9 22 MISSING (IN CYP81B1S).
 FT CONFLICT 40 40 MISSING (IN CA04116).
 FT CONFLICT 51 51 Y->S (IN CA04116).
 FT CONFLICT 78 78 P->Q (IN CA04116).
 SQ SEQUENCE 520 AA; 58914 MW; 38108E379BDB35C8 CRC64;

Query Match 70.0%; Score 35; DB 10; Length 520;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 IGHLYYL 10
 :|||:|
 DB 55 IGHLYYL 61

RESULT 10
 ID 09RM20 PRELIMINARY; PRT: 531 AA.
 AC 09RM20;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE PKO2-42.
 GN Bacillus anthracis.
 OS Bacillus anthracis.
 OG Bacillid PKO2.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 NCBI_TaxID=1392;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Okinaka R.T., Cloud K., Hampton O., Hill K.K., Keim P., Lanke G.,
 RA Kumano S., Maiter D., Martinez Y., Svensson R., Tatum L.R.,
 RA Brown A.E., Jackson P.J.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF188935; AAF13647.1; -
 DR InterPro: IPR001119; -
 DR InterPro: IPR002508; -
 DR InterPro: IPR00395; SLH; 3.
 DR Pfam: PF00395; SLH; 3.
 DR Pfam: PF01520; Amidase_3; 1.
 KW Plasmid.
 SQ SEQUENCE 531 AA; 58864 MW; 9A172DC4ED05CA78 CRC64;

Query Match 70.0%; Score 35; DB 2; Length 531;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 VVRIGHLYL 10
 :|||:|
 DB 467 VKIGNLYL 475

RESULT 11
 ID 006769 PRELIMINARY; PRT: 637 AA.
 AC 006769;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE HYPOHETICAL 69.5 KDA PROTEIN.
 GN RV0669C OR MT1376.05.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
 NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;

```

RX MEDLINE-98295987; PubMed-9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekela F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Deltell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagers K., Bellwell T., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter K., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
DR EMBL: Z95972; CA00388.1; -.
DR Tuberculin; P0669C; -.
KW Hypothetical protein.
SQ
SEQUENCE 637 AA; 69489 MW; 5E2D915AB1E4FDB7 CRC64;

Query Match
Best Local Similarity 70.0%; Score 35; DB 2; Length 637;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VWRIGLHYTL 10
DB 403 LVRIQLHYTL 412

RESULT 12
Q9UDU6 PRELIMINARY; PRT; 723 AA.
ID Q9UDU6;
AC Q9UDU6;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE HEPATOCYTE GROWTH FACTOR.
DE HGF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RA Courtney L., Elliot G., Angell S.;
RT "The sequence of Homo sapiens PAC clone RP5-1098B1.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Waterston R.H.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO SERINE PROTEASES, TRYPsin FAMILY.
CC -1- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR EMBL: AC004960; AAC71655.1; -.
DR HSSP; P14210; 1BHT.
DR InterPro; IPR000001; -.
DR InterPro; IPR000327; -.
DR InterPro; IPR001254; -.
DR InterPro; IPR001314; -.
DR InterPro; IPR003014; -.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00051; kringie; 4.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000583; -. 1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR SMART; SM00020; Tryp_Spc; 1.
KW Hydrolyase; Serine protease.
SQ
SEQUENCE 723 AA; 82602 MW; 627B1EF99FAD931B CRC64;

```

```

Query Match
Best Local Similarity 70.0%; Score 35; DB 4; Length 723;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 VWRIGLHYTL 10
DB 623 LVRIQLHYTL 632

RESULT 13
Q9JLV7 PRELIMINARY; PRT; 808 AA.
ID Q9JLV7;
AC Q9JLV7;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE F2265.28.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RP SEQUENCE FROM N.A.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bel O., Chin C., Chlou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F2265 from chromosome
RT 1.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bel O., Chin C., Chlou J., Choi E.,
RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC022464; AAF79556.1; -.
SQ
SEQUENCE 808 AA; 90243 MW; 628C800CF1AE04F CRC64;

Query Match
Best Local Similarity 70.0%; Score 35; DB 10; Length 808;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VWRIGLHYTL 10
DB 526 LVRIQLHYTL 535

RESULT 14
Q9JLV7 PRELIMINARY; PRT; 1863 AA.
ID Q9JLV7;
AC Q9JLV7;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)

```

DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE TRANSIENT RECEPTOR POTENTIAL-RELATED PROTEIN.
 GN CHAK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Matsushita M.;
 RL Submitted (MAY-1999) to the EMBL/genbank/DBJ databases.
 DR EMBL: AF149013; AAF73131.1; -.
 DR InterPro: IPR002111; -.
 DR InterPro: IPR002153; -.
 DR Pfam: PF02164; Trans_recep; 1.
 KM Receptor.
 SQ SEQUENCE 1863 AA; 212425 MW; B4F0B09BF487EBA CRC64;

Query Match 70.0%; Score 35; DB 11; Length 1863;
 Best Local Similarity 50.0%; Pred. No. 4.9e+02;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 VVRIGHLXYL 10
 Db 1636 ILKSGHLIYI 1645

RESULT 15
 049122
 ID 049122 PRELIMINARY; PRT; 139 AA.
 AC 049122;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE MAV GENE CLUSTER, METHYLAMINE DEHYDROGENASE LARGE AND SMALL SUBUNITS,
 DE AND AMICANIN, (MAUFEBDAGLGM) GENES, COMPLETE CDS (FRAGMENT).
 OS Methylobacterium extorquens.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 OX NCBI_TaxID=408;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AM1;
 RX MEDLINE=94292425; PubMed=8021187;
 RA Chistoserdov A.Y., Chistoserdova L.V., McIntire W.S., Lidstrom M.E.;
 RT "Genetic organization of the mau gene cluster in Methylobacterium
 RT extorquens AM1: complete nucleotide sequence and generation and
 RT characteristics of mau mutants."
 RL J. Bacteriol. 176:4052-4065(1994).
 DR EMBL: L26406; ABB46931.1; -.
 DR InterPro: IPR001387; -.
 DR Pfam: PF01381; HTH_3; 1.
 DR SMART: SM00530; HTH_XRE; 1.
 FT NON_TER 139
 SQ SEQUENCE 139 AA; 14906 MW; 5CEFA626D3183B1 CRC64;

Query Match 68.0%; Score 34; DB 2; Length 139;
 Best Local Similarity 55.6%; Pred. No. 50;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VRIIGLYYL 10
 Db 104 IKVGHLYEL 112

Search completed: June 20, 2001, 14:10:23
 Job time: 345 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2001, 14:04:01 ; Search time 41.58 Seconds

(without alignments)
13.748 Million cell updates/sec

Title: US-09-692-401-5
Perfect score: 50
Sequence: 1 VVRIGHLXIL 10

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 390729 seqs, 5716335 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	84.0	9	15	R50283
2	42	84.0	9	17	R99345
3	42	84.0	9	20	V46062
4	42	84.0	10	18	W23043
5	37	74.0	10	20	V47343
6	36	72.0	363	21	G46573
7	36	72.0	365	21	G13456
8	36	72.0	365	21	G52966
9	36	72.0	369	21	G31496
10	36	72.0	372	21	G26363
11	36	72.0	372	21	G46572

12	36	72.0	378	21	G31495	Arabidopsis thalia
13	36	72.0	394	21	G26362	Arabidopsis thalia
14	36	72.0	421	21	G26361	Arabidopsis thalia
15	36	72.0	421	21	G46571	Arabidopsis thalia
16	36	72.0	423	21	G31455	Arabidopsis thalia
17	36	72.0	423	21	G52965	Arabidopsis thalia
18	36	72.0	427	21	G31494	Arabidopsis thalia
19	35	70.0	189	13	R25692	Arabidopsis thalia
20	35	70.0	320	20	W98133	Exons XV to XVII
21	35	70.0	505	20	W98133	Melanoma protein X
22	35	70.0	621	20	V34971	Jerusalem artichok
23	35	70.0	697	21	V94855	Amino acid sequenc
24	35	70.0	697	21	V59030	Hepatocyte growth
25	35	70.0	697	22	B45838	Sequence of a pept
26	35	70.0	723	11	R07144	Nucleic acid trans
27	35	70.0	723	12	R14307	Tumour cytotoxic f
28	35	70.0	723	12	R15624	Plasminogen-like g
29	35	70.0	723	13	R21142	Human leukocyte-de
30	35	70.0	723	13	R25677	Human TCF-II. Hom
31	35	70.0	723	13	R29819	Recombinant human
32	35	70.0	723	13	R57026	Human wild-type tu
33	35	70.0	723	15	R57027	Human modified tum
34	35	70.0	723	15	R57028	Human modified tum
35	35	70.0	723	16	R82685	Tumour cytotoxic f
36	35	70.0	723	17	R99688	TCF mutant having
37	35	70.0	723	17	R99689	TCF mutant having
38	35	70.0	723	19	W76690	Human plasminogen-
39	35	70.0	723	19	W59923	Human leukocyte-de
40	35	70.0	727	12	R10556	Hepatic parenchyma
41	35	70.0	728	12	R12792	Human hepatocyte g
42	35	70.0	728	12	R14243	Human hepatocyte g
43	35	70.0	728	12	R15623	Human leukocyte-de
44	35	70.0	728	13	R20005	Human hepatocyte g
45	35	70.0	728	13	R20100	Contains alpha- an

ALIGNMENTS

RESULT 1	
R50283	
ID R50283	standard. Protein; 9 AA.
XX	
AC R50283;	
XX	
DT 26-SEP-1994	(first entry)
XX	
DE	MAGE-21 nonapeptide.
XX	
KW	MAGE; nonapeptide; cancer; melanoma; breast cancer; HLA;
KW	histocompatibility; human leucocyte antigen; probe; treatment;
KW	therapy; vaccine.
XX	
OS	Synthetic.
XX	
PN	W09405304-A.
XX	
PD	17-MAR-1994.
XX	
PF	30-AUG-1993; 93WO-0508157.
XX	
PR	31-AUG-1992; 92US-0938334.
PR	26-MAR-1993; 93US-0037230.
PR	07-JUN-1993; 93US-0073103.
XX	
RA	(LUDW-) LUDWIG INST CANCER RES.
XX	
PI	Boon-falleur T, De Plaen E, Lurquin C, Traversari C;
PI	Van Derbruggen P;
XX	
DR	WPI; 1994-100844/12.
DR	N-PSDB; Q44753.
XX	

PR New nona:peptide derived from tumour rejection antigen precursor
PT - presented by HLA-A1 cancer cells, for use in diagnosis or
PT therapy of esp. melanoma and breast cancer.
PS Disclosure: Page 20; 33pp; English.
XX
XX An isolated nonapeptide having the amino acid sequence Glu-Val-Asp-
CC Pro-Ile-Gly-His-Leu-Tyr is derived from the tumour rejection antigen
CC precursor encoded by the MAGE-3 gene and presented by HLA-A1. The
CC nonapeptide can be used in a vaccine to treat a cancerous condition
CC involving HLA-A1 subtype cancerous cells. The nucleic acid encoding
CC the nonapeptide can be used as a probe to identify tumour cells.
CC This sequence is homologous to the peptide described and is encoded
CC by the MAGE-21 gene.
XX
SQ Sequence 9 AA;

Query Match 84.0%; Score 42; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVRIGHLY 8
|||||||
Db 2 vvrighly 9

RESULT 2
R99345
ID R99345 standard; Protein; 9 AA.
XX
AC R99345;

DT 22-APR-1997 (first entry)
XX

DE MAGE-21 nonapeptide.
XX

KM HLA binding peptide; cell lysis; cytolytic T cell; MAGE family; human;
KM tumour rejection antigen precursor; TRA; MAGE-1; tumour; cancer cell;
KM antibody; melanoma; universal effector cell; vaccine; breast cancer; CTL;
KM therapy.
XX

OS Homo sapiens.
XX

PN WO9626214-A1.
XX

PD 29-AUG-1996.
XX

PF 01-FEB-1996; 96WO-US01489.
XX

PR 23-FEB-1995; 95US-0393273.
XX

PA (LUDWIG INST CANCER RES.
XX

PI Boon-Falleur T, De Plaen E, Gaugler B, Lurquin C;
PI Romero P, Traversari C, Van Den Eynde B, Van Der Bruggen P;
XX

DR WPI: 1996-402317/40.
XX

DR N-PSDB: T35410.
XX

PT New nona:peptide(s) that bind to HLA molecule(s) and induce lysis -
PT by specific cytolytic T cells, for diagnosis and treatment of
PT tumours and to expand T cells in vitro.
XX

PS Example 4; Fig 4; 41pp; English.
XX

CC R99343-R99350 represent MAGE nonapeptides, based on the tumour rejection
CC antigen region of the full length MAGE sequences. These peptides were
CC used to design the nonapeptides of the invention. (see R99337-R99342),
CC which bind to a HLA molecule on a cell, and provoke lysis by cytolytic T
CC cells (CTLs) specific for a complex of the HLA molecule and nonapeptide.
CC The nonapeptides can be used diagnostically to identify tumours
CC expressing a particular HLA molecule, or to identify cancer cells. The

CC peptides can also be used therapeutically, to induce a CTL response to
CC tumours (where the peptides are optionally coupled to tumour-specific
CC antibodies), or to induce a response by CTLs that are otherwise inactive.
CC The peptide sequences may also be used to expand specific CTLs in vitro
CC for later return to the patient, such as for treating melanoma. Tumour
CC cells can be identified by using DNA encoding the nonapeptides as probes.
CC Non-human cells transformed with the HLA-A1 gene and a DNA sequence
CC encoding one of the peptides, can be used to generate CTLs, or to detect
CC the presence of CTLs in human samples. The non-human transformed cells,
CC when polytransformed, are universal effector cells, and can be used in
CC vaccines, or for treating melanoma or breast cancer.
XX
SQ Sequence 9 AA;

Query Match 84.0%; Score 42; DB 17; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VVRIGHLY 8
|||||||
Db 2 vvrighly 9

RESULT 3
Y46062
ID Y46062 standard; Peptide; 9 AA.
XX
AC Y46062;

DT 01-DEC-1999 (first entry)
XX

DE Immunogenic peptide having a human leukocyte antigen binding motif #673.
XX

KM Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KM immune response; T cell activation; major histocompatibility complex;
KM cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KM prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KM vaccine; immunisation.
XX

OS Synthetic.
XX

PN WO9945954-A1.
XX

PD 16-SEP-1999.
XX

PF 13-MAR-1998; 98WO-US05039.
XX

PR 13-MAR-1998; 98WO-US05039.
XX

PA (EPIM-) EPIMUNE INC.
XX

PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
PI WPI: 1999-551214/46.
XX

DR
XX

PT New immunogenic peptides with HLA binding motif, useful in treatment
PT and diagnosis of cancers and viral diseases -
XX

PS Claim 1; Page 54; 150pp; English.
XX

CC Y45390 to Y48214 represent specifically claimed immunogenic peptides
CC having a human major histocompatibility complex (MHC) Class I (also
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
CC response against the antigen from which the peptide is derived.
CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
CC normally induced by an antigen in the form of a peptide fragment bound
CC to a HLA molecule, rather than the intact foreign antigen itself, and
CC are particularly important in tumour rejection and in fighting viral
CC infections. The peptides are therefore useful therapeutically to treat

CC or prevent viral infections and cancers in mammals (especially humans)
 CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
 CC They can be administered as vaccines to elicit an immune response in
 CC individuals susceptible or otherwise at risk of viral infection or
 CC cancer, or used to treat chronic or acute conditions. They are also
 CC useful diagnostically, and can be used to induce a cytotoxic T cell
 CC response, by contacting a cytotoxic T cell with the peptide e.g. to
 CC produce CTLs ex vivo for infusion back into a patient. The
 CC polynucleotides encoding the immunogenic peptides are also useful
 CC therapeutically and for immunisation as above.
 CC
 SQ Sequence 9 AA:

Query Match 84.0%; Score 42; DB 20; Length 9;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVRIGHLY 8
 |||||
 Db 2 VVRIGHLY 9

RESULT 4

W23043 23043 standard; Peptide: 10 AA.

AC W23043;

XX 25-FEB-1998 (first entry)

DE MAGF-12/HLA-B44 tumour rejection antigen.

XX MAGF-12; tumour rejection antigen precursor; TRAP; HLA-B44;

KW human leukocyte antigen B44; cytotoxic T lymphocyte; cancer;

KM melanoma; therapy; diagnosis; vaccine.

OS Homo sapiens.

PN W09731017-A1.

PD 28-AUG-1997.

XX 05-FEB-1997; 97WO-US01915.

XX 20-FEB-1996; 96US-0602506.

PA (LUDW-) LUDWIG INST CANCER RES.

PI Boon-Falleur T, Coulie P, Herman J, Luescher I;

PI Van Der Bruggen P;

PT Tumour rejection antigens presented by human leukocyte antigen B44

PS Claim 2; Page 51; 74pp; English.

CC This peptide is a tumour rejection antigen presented by a HLA-B44
 CC molecule and derived from a MAGF-12 tumour rejection antigen
 CC precursor (TRAP). Claimed tumour rejection antigens (W23038-43)
 CC are able to bind to HLA-B44 positive cells, making them useful in
 CC identifying cells which present HLA-B44 molecules on their
 CC surfaces for use in the diagnosis and therapy of cellular
 CC abnormalities. The complex of the tumour rejection antigen and HLA
 CC molecule provokes a cytolytic T cell response. The tumour
 CC rejection antigens, or complexes of tumour rejection antigens and
 CC HLA-B44, can be used as vaccines to treat disorders characterised
 CC by expression of the TRAP molecule such as cancer, especially
 CC melanoma. Vaccines can also be prepared from cells which present
 CC the tumour rejection antigen/HLA complexes on their surface, such

CC as non-proliferative cancer cells and non-proliferative
 CC transfectants.
 CC
 SQ Sequence 10 AA;

Query Match 84.0%; Score 42; DB 18; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.016;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VVRIGHLY 8
 |||||
 Db 3 VVRIGHLY 10

RESULT 5

Y47343 Y47343 standard; Peptide: 10 AA.

AC Y47343;

XX 01-DEC-1999 (first entry)

DE Immunogenic peptide having a human leukocyte antigen binding motif #1954.

KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;

KW immune response; T cell activation; major histocompatibility complex;

KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;

KM prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;

OS Synthetic.

PN W09945954-A1.

PD 16-SEP-1999.

XX 13-MAR-1998; 98WO-US05039.

XX 13-MAR-1998; 98WO-US05039.

PA (EPIW-) EPIMUNE INC.

PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;

PI WPI: 1999-551214/46.

PT New immunogenic peptides with HLA binding motif, useful in treatment

PS Claim 1; Page 104; 150pp; English.

CC Y45390 to Y48214 represent specifically claimed immunogenic peptides
 CC having a human major histocompatibility complex (MHC) Class I (also
 CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
 CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes
 CC HLA-A2.1, A1, A2.2 or A2.1 or HLA-B or C) and induce a cytotoxic T cell
 CC response against the antigen from which the peptide is derived.
 CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
 CC normally induced by an antigen in the form of a peptide fragment bound
 CC to a HLA molecule, rather than the intact foreign antigen itself, and
 CC are particularly important in tumour rejection and in fighting viral
 CC infections. The peptides are therefore useful therapeutically to treat
 CC or prevent viral infections and cancers in mammals (especially humans)
 CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
 CC They can be administered as vaccines to elicit an immune response in
 CC individuals susceptible or otherwise at risk of viral infection or
 CC cancer, or used to treat chronic or acute conditions. They are also
 CC useful diagnostically, and can be used to induce a cytotoxic T cell
 CC response, by contacting a cytotoxic T cell with the peptide e.g. to
 CC produce CTLs ex vivo for infusion back into a patient. The
 CC polynucleotides encoding the immunogenic peptides are also useful

CC therapeutically and for immunisation as above.
SQ Sequence 10 AA:

Query Match 74.0%; Score 37; DB 20; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.17;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VVRIGHLXIL 10
|||
Db 1 vvpishlyil 10

RESULT 6
G46573
ID G46573 standard; Protein; 363 AA.

AC G46573;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 58606.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.

XX Arabidopsis thaliana.

XX Ep1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-011825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125768.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145918.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0144793.
PR 09-AUG-1999; 99US-0144793.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156438.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 72.0%; Score 36; DB 21; Length 363;
Best Local Similarity 77.8%; Pred. No. 15;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 VRIGHTLY 9
||:|||||
Db 8 vymghlyl 16
RESULT 7
ID G13456 standard; Protein; 365 AA.
XX
AC G13456;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 12958.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134259.
PR 14-MAY-1999; 99US-0134261.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134378.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154019.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161820.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 72.0%; Score 36; DB 21; Length 365;
Best Local Similarity 66.7%; Pred. No. 15;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VVRIGHLXI 9

Db :|:|||||
8 lvy|ghlyl 16

RESULT 8
ID G52966 standard; Protein; 365 AA.
XX
AC G52966;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 67386.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130649.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0143542.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
XX 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157573.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 72.0%; Score 36; DB 21; Length 365;
Best Local Similarity 66.7%; Pred. NO. 15;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VVRIGHLYI 9
: | : | | | | |
Db 8 lvyighlyl 16

RESULT 9
G31496
ID G31496 standard; Protein; 369 AA.
XX
AC G31496;
XX
DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 37833.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132408.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137232.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.

```
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142972.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
```

```
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151348.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

```
Query Match 72.0%; Score 36; DB 21; Length 369;
Best Local Similarity 77.8%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 VVRIGHLTY 9
||:|||||
DB 8 vvygmghlyl 16
```

```
RESULT 10
G26363
ID G26363 standard; Protein; 372 AA.
```

```
XX G26363;
```

```
DT 17-OCT-2000 (first entry)
```

```
DE Arabidopsis thaliana protein fragment SEQ ID NO: 30791.
```

```
XX Protein identification: signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
```

OS	Arabidopsis thaliana.	
XX	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
PF	25-FEB-2000; 2000EP-0301439.	
XX		
PR	25-FEB-1999; 99US-0121825.	PR 30-JUN-1999; 99US-0141287.
PR	05-MAR-1999; 99US-0123160.	PR 01-JUL-1999; 99US-0141842.
PR	09-MAR-1999; 99US-0123388.	PR 01-JUL-1999; 99US-0142154.
PR	23-MAR-1999; 99US-0125788.	PR 02-JUL-1999; 99US-0142055.
PR	25-MAR-1999; 99US-0126264.	PR 06-JUL-1999; 99US-0142390.
PR	29-MAR-1999; 99US-0126785.	PR 08-JUL-1999; 99US-0142803.
PR	01-APR-1999; 99US-0127462.	PR 09-JUL-1999; 99US-0142920.
PR	06-APR-1999; 99US-0128234.	PR 12-JUL-1999; 99US-0142977.
PR	08-APR-1999; 99US-0128714.	PR 13-JUL-1999; 99US-0143542.
PR	16-APR-1999; 99US-0129845.	PR 14-JUL-1999; 99US-0143624.
PR	19-APR-1999; 99US-0130077.	PR 15-JUL-1999; 99US-0144005.
PR	21-APR-1999; 99US-0130449.	PR 16-JUL-1999; 99US-0144085.
PR	23-APR-1999; 99US-0130510.	PR 16-JUL-1999; 99US-0144086.
PR	23-APR-1999; 99US-0130891.	PR 19-JUL-1999; 99US-0144325.
PR	28-APR-1999; 99US-0131449.	PR 19-JUL-1999; 99US-0144331.
PR	30-APR-1999; 99US-0132048.	PR 19-JUL-1999; 99US-0144332.
PR	30-APR-1999; 99US-0132407.	PR 19-JUL-1999; 99US-0144333.
PR	04-MAY-1999; 99US-0132484.	PR 19-JUL-1999; 99US-0144334.
PR	05-MAY-1999; 99US-0132485.	PR 19-JUL-1999; 99US-0144335.
PR	06-MAY-1999; 99US-0132486.	PR 20-JUL-1999; 99US-0144352.
PR	06-MAY-1999; 99US-0132487.	PR 20-JUL-1999; 99US-0144632.
PR	07-MAY-1999; 99US-0132863.	PR 20-JUL-1999; 99US-0144684.
PR	11-MAY-1999; 99US-0134256.	PR 21-JUL-1999; 99US-0145086.
PR	14-MAY-1999; 99US-0134218.	PR 21-JUL-1999; 99US-0145088.
PR	14-MAY-1999; 99US-0134221.	PR 21-JUL-1999; 99US-0145088.
PR	14-MAY-1999; 99US-0134370.	PR 22-JUL-1999; 99US-0145085.
PR	18-MAY-1999; 99US-0134768.	PR 22-JUL-1999; 99US-0145087.
PR	19-MAY-1999; 99US-0134941.	PR 22-JUL-1999; 99US-0145089.
PR	20-MAY-1999; 99US-0135124.	PR 22-JUL-1999; 99US-0145192.
PR	21-MAY-1999; 99US-0135353.	PR 23-JUL-1999; 99US-0145145.
PR	24-MAY-1999; 99US-0135629.	PR 23-JUL-1999; 99US-0145218.
PR	25-MAY-1999; 99US-0136021.	PR 23-JUL-1999; 99US-0145224.
PR	27-MAY-1999; 99US-0136392.	PR 26-JUL-1999; 99US-0145276.
PR	28-MAY-1999; 99US-0136782.	PR 27-JUL-1999; 99US-0145913.
PR	01-JUN-1999; 99US-0137232.	PR 27-JUL-1999; 99US-0145918.
PR	03-JUN-1999; 99US-0137528.	PR 27-JUL-1999; 99US-0145819.
PR	04-JUN-1999; 99US-0137502.	PR 28-JUL-1999; 99US-0145951.
PR	07-JUN-1999; 99US-0137724.	PR 02-AUG-1999; 99US-0146386.
PR	08-JUN-1999; 99US-0138094.	PR 02-AUG-1999; 99US-0146388.
PR	10-JUN-1999; 99US-0138540.	PR 02-AUG-1999; 99US-0146389.
PR	10-JUN-1999; 99US-0138847.	PR 03-AUG-1999; 99US-0147038.
PR	14-JUN-1999; 99US-0139119.	PR 04-AUG-1999; 99US-0147204.
PR	16-JUN-1999; 99US-0139452.	PR 04-AUG-1999; 99US-0147302.
PR	16-JUN-1999; 99US-0139453.	PR 05-AUG-1999; 99US-0147192.
PR	17-JUN-1999; 99US-0139452.	PR 05-AUG-1999; 99US-0147260.
PR	18-JUN-1999; 99US-0139454.	PR 06-AUG-1999; 99US-0147303.
PR	18-JUN-1999; 99US-0139455.	PR 06-AUG-1999; 99US-0147416.
PR	18-JUN-1999; 99US-0139456.	PR 09-AUG-1999; 99US-0147493.
PR	18-JUN-1999; 99US-0139457.	PR 10-AUG-1999; 99US-0147935.
PR	18-JUN-1999; 99US-0139458.	PR 10-AUG-1999; 99US-0148171.
PR	18-JUN-1999; 99US-0139459.	PR 11-AUG-1999; 99US-0148341.
PR	18-JUN-1999; 99US-0139460.	PR 12-AUG-1999; 99US-0148341.
PR	18-JUN-1999; 99US-0139461.	PR 13-AUG-1999; 99US-0148565.
PR	18-JUN-1999; 99US-0139462.	PR 13-AUG-1999; 99US-0148684.
PR	18-JUN-1999; 99US-0139463.	PR 16-AUG-1999; 99US-0149368.
PR	18-JUN-1999; 99US-0139750.	PR 17-AUG-1999; 99US-0149175.
PR	18-JUN-1999; 99US-0139763.	PR 18-AUG-1999; 99US-0149426.
PR	21-JUN-1999; 99US-0139817.	PR 20-AUG-1999; 99US-0149722.
PR	22-JUN-1999; 99US-0139829.	PR 20-AUG-1999; 99US-0149723.
PR	23-JUN-1999; 99US-0140353.	PR 20-AUG-1999; 99US-0149929.
PR	23-JUN-1999; 99US-0140354.	PR 23-AUG-1999; 99US-0149902.
PR	*24-JUN-1999; 99US-0140695.	PR 23-AUG-1999; 99US-0149930.
PR	28-JUN-1999; 99US-0140823.	PR 25-AUG-1999; 99US-0150566.
PR	29-JUN-1999; 99US-0140991.	PR 26-AUG-1999; 99US-0150884.

PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0155658.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159284.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 23-OCT-1999; 9905-0161404.
PR 23-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161923.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

Query Match 72.0%; Score 36; DB 21; Length 372;
Best Local Similarity 77.8%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Caps 0;

OY 1 VVRIGHLXI 9
DB 17 vvygmghlyi 25

RESULN 11
G46572 G46572 standard; Protein; 372 AA.

AC G46572:

DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 58605.

KW protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

XX Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 9905-0121825.
PR 05-MAR-1999; 9905-0123180.
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0129845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130449.
PR 23-APR-1999; 9905-0130510.
PR 23-APR-1999; 9905-0130891.
PR 28-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 30-APR-1999; 9905-0132407.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 07-MAY-1999; 9905-0132487.
PR 11-MAY-1999; 9905-0132863.
PR 14-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135253.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137528.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139453.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139753.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.

PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145112.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147182.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154078.
PR 16-SEP-1999; 99US-0154079.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159568.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 72.0%; Score 36; DB 21; Length 372;
Best Local Similarity 77.8%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VVRIGHLYI 9
||:|||||
Db 17 vvygmhlyl 25
RESULT 12
G31495
ID G31495 standard; Protein; 378 AA.
AC G31495;
XX 17-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana protein fragment SEQ ID NO: 37832.
DE
XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX Arabidopsis thaliana.
PN EP1033405-A2.
PD
XX 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130049.
PR 23-APR-1999; 990S-0130510.
PR 28-APR-1999; 990S-0130891.
PR 30-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 04-MAY-1999; 990S-0132407.
PR 05-MAY-1999; 990S-0132484.
PR 06-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 07-MAY-1999; 990S-0132487.
PR 11-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142877.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.

PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144684.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147392.
PR 05-AUG-1999; 990S-0147360.
PR 06-AUG-1999; 990S-0147703.
PR 06-AUG-1999; 990S-0147816.
PR 09-AUG-1999; 990S-0147893.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158369.

PR	23-APR-1999	99US-0130891
PR	28-APR-1999	99US-0131449
PR	30-APR-1999	99US-0132048
PR	04-MAY-1999	99US-0132407
PR	05-MAY-1999	99US-0132484
PR	06-MAY-1999	99US-0132485
PR	06-MAY-1999	99US-0132486
PR	07-MAY-1999	99US-0132487
PR	11-MAY-1999	99US-0132863
PR	11-MAY-1999	99US-0134256
PR	11-MAY-1999	99US-0134218
PR	14-MAY-1999	99US-0134219
PR	14-MAY-1999	99US-0134221
PR	14-MAY-1999	99US-0134370
PR	18-MAY-1999	99US-0134741
PR	19-MAY-1999	99US-0134748
PR	20-MAY-1999	99US-0135124
PR	21-MAY-1999	99US-0135553
PR	24-MAY-1999	99US-0135629
PR	25-MAY-1999	99US-0136021
PR	27-MAY-1999	99US-0136592
PR	28-MAY-1999	99US-0137222
PR	01-JUN-1999	99US-0137682
PR	03-JUN-1999	99US-0137528
PR	04-JUN-1999	99US-0137502
PR	07-JUN-1999	99US-0137724
PR	08-JUN-1999	99US-0138094
PR	10-JUN-1999	99US-0138540

PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.

PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.

PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.

PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.

22-JUN-1999; 99US-0139899.

PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140861.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0142154.

PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.

PR 13-JUL-1999; 99US-0143542.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144086.
 PR 10-TRF-1000; 00US-0144305

PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999.	99US-0144332

PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144333.

PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.

EN	20 JUL 1999;	99US-0144632.
PR	20-JUL-1999;	

```
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145222.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150889.
PR 27-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158233.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159299.
PR 13-OCT-1999; 99US-0159299.
PR 13-OCT-1999; 99US-0159299.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
```

```
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 72.0%; Score 36; DB 21; Length 394.
Best Local Similarity 77.8%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VVRIGLDYI 9
  ||:|||||
Db 39 vvygmghlyl 47

RESULT 14
ID G26361 standard; Protein: 421 AA.
XX
AC G26361;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SQ ID NO: 30789.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-AZ.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
```

PR 06-MAY-1999; 99US-0133267.
PR 07-MAY-1999; 99US-0132663.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139482.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 26-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0144297.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149729.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 11-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.

PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 72.0%; Score 36; DB 21; Length 421;
Best Local Similarity 77.8%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VVRGHLXI 9
11:111111
Db 66 vvygmghlyl 74

RESULT 15
G46571
ID G46571 standard; Protein: 421 AA.

AC G46571;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 58604.

KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PE 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0133484.

PR 05-MAY-1999; 99US-0133485.

PR 06-MAY-1999; 99US-0133486.

PR 07-MAY-1999; 99US-0133487.

PR 11-MAY-1999; 99US-0134266.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137702.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0143277.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147382.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150366.
PR 23-AUG-1999; 99US-0150366.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159684.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 72.0%; Score 36; DB 21; Length 421;
Best Local Similarity 77.8%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VVRIGHLYI 9
||:|||||
Db 66 vvygmhlyi 74

Search completed: June 20, 2001, 14:04:02
Job time: 49 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2001, 14:03:13 ; Search time 28.81 Seconds
(without alignments)
23,796 Million cell updates/sec

Title: US-09-692-401-4

Perfect score: 46

Sequence: 1 VRIGHLYIL 9

Scoring table:

BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	314	2	melanoma antigen M
2	35	76.1	233	2	protein F23N19.5
3	35	76.1	234	2	probable membrane
4	35	76.1	305	2	hypothetical prote
5	35	76.1	505	2	cytochrome P450 (E
6	35	76.1	613	2	Gcpe protein limpo
7	35	76.1	613	2	hypothetical prote
8	34	73.9	637	2	hepatocytic growth
9	34	73.9	728	1	hepatocytic growth
10	34	73.9	728	1	hepatocytic growth
11	34	73.9	728	1	hepatocytic growth
12	34	73.9	1244	2	hypothetical prote
13	34	73.9	1461	2	probable retroel
14	34	73.9	313	2	calcium-binding pr
15	33	71.7	314	2	melanoma antigen M
16	33	71.7	314	2	melanoma antigen M
17	33	71.7	314	2	melanoma antigen M
18	33	71.7	337	2	probable ethylene
19	33	71.7	423	2	CDP-diacylglycerol
20	33	71.7	423	2	CDP-diacylglycerol
21	32	69.6	619	1	glucokinase regula
22	32	69.6	256	2	tonoplast intrinsi
23	32	69.6	325	2	hypothetical prote
24	32	69.6	350	2	hypothetical prote
25	32	69.6	354	2	N-acetylglucosamin
26	32	69.6	368	2	hypothetical prote
27	32	69.6	415	2	proximal sequence
28	32	69.6	427	2	heterogeneous nucl
29	32	69.6	449	2	CDP-diacylglycerol
					hypothetical prote

30	32	69.6	449	2	139358	heterogeneous nucl
31	32	69.6	889	2	H96606	hypothetical prote
32	32	69.6	2241	2	S09811	hypothetical prote
33	31	67.4	101	2	E82849	conserved hypotnet
34	31	67.4	276	2	T12552	hypothetical prote
35	31	67.4	295	2	H83429	hypothetical prote
36	31	67.4	346	2	T48496	hypothetical prote
37	31	67.4	358	1	JX0286	membrane protein -
38	31	67.4	379	1	JX0286	3-Isopropylmalate
39	31	67.4	385	2	C70124	conserved hypotnet
40	31	67.4	424	2	T07366	hypothetical prote
41	31	67.4	438	2	T04800	probable phosphati
42	31	67.4	444	1	E69130	CDP-diacylglycerol
43	31	67.4	451	2	T49095	histidine--tRNA 11
44	31	67.4	465	2	C70594	beta-ketocacyl-CoA
45	31	67.4	484	2	C84955	probable phosphoma
						UDP-N-acetylmutrama

ALIGNMENTS

RESULT 1
154519
melanoma antigen MAGE-12 - human
N:Alternate names: MAGE 21 protein
C:Species: Homo sapiens (man)
C>Date: 07-Jun-1996 #sequence-revision 07-Jun-1996 #text-change 18-Feb-2000
C:Accession: I54519; J02362; PH1295
R:De Smet, C.; Lurquin, C.; van der Bruggen, P.; De Plaen, E.; Brasseur, F.; Boon, T.
Immunogenetics 39, 121-129, 1994
A:Title: Sequence and expression pattern of the human MAGE2 gene.
A:Reference number: I54519; MUID:94102805
A:Accession: I54519
A>Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-314 <DS>
A:Cross-references: GB:I18877; NID:9499345; PIDN:AAA19023.1; PID:9499346
R:Dirig, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.
Biochem. Biophys. Res. Commun. 202, 549-555, 1994
A:Title: Cloning and analysis of MAGE-1-related genes.
A:Reference number: J02358; MUID:94311935
A:Accession: J02362
A:Molecule type: mRNA
A:Residues: 1-9, 'S', 11-186, 'D', 188-299, 'S', 301-314 <DIN>
A:Experimental source: melanoma cell line DM150; MAGE-12f
R:Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin, C.; Chomez, P.; Van, P.
J. Exp. Med. 176, 1453-1457, 1992
A:Title: A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytoly
A:Reference number: PH1294; MUID:93018875
A:Accession: PH1295
A:Molecule type: DNA
A:Residues: 168-176 <TRA>
A:Experimental source: MAGE-21
C:Genetics:
A:Gene: GDB:MAGEA12; MAGE12; MAGE-12f
A:Cross-references: GDB:311129
A:Map position: Xq28-Xq28
C:Superfamily: tumor associated protein MAGE
F:168-176/Region: HLA-A1 binding #status predicted

Query Match 100.0%; Score 46; DB 2; Length 314;
Best local Similarity 100.0%; Pred. No. 0.085;
Matches 9; Conservative 0; Indels 0; Gaps 0;

OY 1 VRIGHLYIL 9
|||||||
Db 170 VRIGHLYIL 178

RESULT 2
B96652
protein F23N19.5 [Imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: B96552
 R:Theologos, A.; Ecker, J.R.; Palm, C.D.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzall,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: B96552
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-233 <STO>
 A:Cross-references: GB:AE005173; NID:g6630448; PIDN:AAFI9536.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F23N19.5
 A:Map position: 1

Query Match 76.1%; Score 35; DB 2; Length 233;
 Best Local Similarity 85.7%; Pred. No. 11;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRIGHLY 7
 Db 177 VRIGHVY 183

RESULT 3
 T37141
 probable membrane protein - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T37141
 R:Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rastread, M.A.
 submitted to the EMBL Data Library, August 1999
 A:Reference number: Z21622
 A:Accession: T37141
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-234 <HAR>
 A:Cross-references: EMBL:AI109972; PIDN:CAB53266.1; GSPDB:GN00070; SCOEDB:SCJ9A.05c
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SCJ9A.05c

Query Match 76.1%; Score 35; DB 2; Length 234;
 Best Local Similarity 55.6%; Pred. No. 11;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VRIGHLY 9
 Db 148 LEIGHMYL 156

RESULT 4
 I40565
 hypothetical protein 4.60 - Bacillus subtilis plasmid pTA1060
 C:Species: Bacillus subtilis
 C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 24-Nov-1999
 C:Accession: I40565
 R:Meijer, W.J.; Venema, G.; Bron, S.
 Nucleic Acids Res. 23, 612-619, 1995
 A:Title: Characterization of single strand origins of cryptic rolling-circle plasmids fr
 A:Reference number: I40549; MUID:95206941
 A:Accession: I40565
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-305 <RES>
 A:Cross-references: EMBL:U32280; NID:g1049123; PIDN:AA04422.1; PID:g1049128
 A:Experimental source: plasmid pTA1060
 C:Genetics:
 A:Gene: plasmid
 C:Superfamily: Bacillus subtilis plasmid pTA1060 hypothetical protein 4.60

Query Match 76.1%; Score 35; DB 2; Length 305;
 Best Local Similarity 66.7%; Pred. No. 15;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VRIGHLY 9
 Db 197 VLIHMYVL 205

RESULT 5
 T10896
 cytochrome P450 (EC 1.14.-.-) 81B1C - Jerusalem artichoke
 C:Species: Helianthus tuberosus (Jerusalem artichoke)
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Feb-2001
 C:Accession: T10896
 R:Chello-Hurtado, F.; Batard, Y.; Salaun, J.; Durst, F.; Pinot, F.; Werck-Reichhart,
 J. Biol. Chem. 273, 7260-7267, 1998
 A:Title: Cloning, expression in yeast and functional characterization of CYP81B1, a p
 A:Reference number: Z17204; MUID:98184826
 A:Accession: T10896
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-505 <CAB>
 A:Cross-references: EMBL:AJ000477; NID:g3059128; PIDN:CA04116.1; PID:g3059129
 A:Experimental source: cv. blanc commun
 C:Genetics:
 A:Gene: CYP81B1C
 C:Function:
 A:Description: specifically catalyzes the hydroxylation of medium chain saturated fat
 C:Species: human cytochrome P450 CYP2D6; cytochrome P450 homology
 C:Key words: chromoprotein; fatty acid metabolism; heme; iron; metalloprotein; oxidore
 F:301-653/Domain: cytochrome P450 homology <P45>
 F:441/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 76.1%; Score 35; DB 2; Length 505;
 Best Local Similarity 85.7%; Pred. No. 26;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 IGHLYL 9
 Db 40 IGHLYL 46

RESULT 6
 C86537
 Gcpe protein [imported] - Chlamydomonas reinhardtii (strain J138)
 C:Species: Chlamydomonas reinhardtii, Chlamydia pneumoniae
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
 C:Accession: C86537
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A:Reference number: A86491; MUID:20330349
 A:Accession: C86537
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-613 <STO>
 A:Cross-references: GB:BA000008; NID:g6978745; PIDN:BA98581.1; GSPDB:GN00142
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: gcpe

Query Match 76.1%; Score 35; DB 2; Length 613;
 Best Local Similarity 87.5%; Pred. No. 32;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRIGHLYI 8
 |||||
 DB 18 VRIGNLYI 25

RESULT 7
 E72087
 gcpE protein CP0383 [Imported] - Chlamydia pneumoniae (strains CWL029 and AR39)
 C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
 C:Accession: E72087; DB1582
 R:Kilman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
 Nature Genet. 21, 385-389, 1999
 A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
 A:Reference number: A72000; MUID:99206606
 A:Accession: E72087
 A:Molecule type: DNA
 A:Residues: 1-613 <ARN>
 A:Cross-references: GB:AE001621; GB:AE001363; NID:g4376641; PIDN:ADB18517.1; PID:g437665
 A:Experimental source: strain CWL029
 R:Read, T.D.; Brumham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
 C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
 Nucleic Acids Res. 28, 1397-1406, 2000
 A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.
 A:Reference number: AB1500; MUID:20150255
 A:Accession: DB1582
 A:Molecule type: DNA
 A:Residues: 1-613 <REA>
 A:Cross-references: GB:AE002200; GB:AE002161; NID:g7189305; PIDN:AAF38230.1; PID:g718930
 A:Experimental source: strain AR39, HL cells
 C:Genetics:
 A:Gene: gcpE; CP0383

Query Match 76.1%; Score 35; DB 2; Length 613;
 Best Local Similarity 87.5%; Pred. No. 32;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRIGHLYI 8
 |||||
 DB 18 VRIGNLYI 25

RESULT 8
 H70535
 hypothetical protein RV0669c - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: H70535
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
 Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellon, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987
 A:Accession: H70535
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-637 <COL>
 A:Cross-references: GB:J295972; GB:AL123456; NID:g3261790; PIDN:CAB09388.1; PID:g2143290
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: RV0669c
 C:Superfamily: Mycobacterium tuberculosis hypothetical protein RV0669c

Query Match 73.9%; Score 34; DB 2; Length 637;
 Best Local Similarity 66.7%; Pred. No. 54;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VRIGHLYI 9
 |||||
 DB 404 VRIGNLYI 412

RESULT 9
 JH0579
 hepatocyte growth factor precursor [validated] - human
 N:Alternate names: hepatoin A; scatter factor
 C:Species: Homo sapiens (man)
 C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 08-Dec-2000
 C:Accession: JH0579; J00333; A41140; B36677; A3512; A39006; PH0114; A37796;
 R:Seki, T.; Hagiya, M.; Shimomishi, M.; Nakamura, T.; Shimizu, S.
 Gene 102, 213-219, 1991
 A:Title: Organization of the human hepatocyte growth factor-encoding gene.
 A:Reference number: JH0579; MUID:91340155
 A:Accession: JH0579
 A:Molecule type: DNA
 A:Residues: 1-728 <SEK>
 A:Cross-references: DDBJ:090318
 A:Note: the authors translated the codon GAA for residue 662 as Gly
 R:Seki, T.; Hagiya, M.; Shimomishi, M.; Nakamura, T.; Shimizu, S.
 Submitted to JIPID, March 1991
 A:Description: Organization of the human hepatocyte growth factor-encoding gene.
 A:Reference number: J00333
 A:Accession: J00333
 A:Molecule type: DNA
 A:Residues: 1-481, 728 <SE2>
 A:Weidner, K.M.; Arakaki, N.; Hartmann, G.; Vandekekerckhove, J.; Weingart, S.; Rieder,
 Proc. Natl. Acad. Sci. U.S.A. 88, 7001-7005, 1991
 A:Title: Evidence for the identity of human scatter factor and human hepatocyte growth
 A:Reference number: A41140; MUID:91334393
 A:Accession: A41140
 A:Molecule type: mRNA
 A:Residues: 1-728 <WEI>
 A:Cross-references: GB:M73239; NID:g337935; PIDN:AAA64239.1; PID:g337936
 R:Seki, T.; Ihara, I.; Sugimura, A.; Shimomishi, M.; Nishizawa, T.; Asami, O.; Hagiya,
 Biochem. Biophys. Res. Commun. 172, 321-327, 1990
 A:Title: Isolation and expression of cDNA for different forms of hepatocyte growth fa
 A:Reference number: A36677; MUID:91025062
 A:Accession: B36677
 A:Molecule type: mRNA
 A:Residues: 1-728 <SE3>
 A:Cross-references: GB:M60718; NID:g184031; PIDN:AAA52648.1; PID:g184032
 A:Accession: A36677
 A:Molecule type: mRNA
 A:Residues: 1-161, 167-728 <SE4>
 A:Cross-references: EMBL:X16323
 A:Experimental source: Leukocyte
 R:Myazawa, K.; Tsubouchi, H.; Naka, D.; Takahashi, K.; Okigaki, M.; Arakaki, N.; Nak
 Biochem. Biophys. Res. Commun. 163, 967-973, 1989
 A:Title: Molecular cloning and sequence analysis of cDNA for human hepatocyte growth
 A:Reference number: A3512; MUID:89392017
 A:Accession: A3512
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-728 <MY>
 A:Cross-references: GB:M29145; NID:g184041; PIDN:AAA62650.1; PID:g306846
 R:Rudin, J.S.; Chan, A.M.W.; Bottaro, D.P.; Burgess, W.H.; Taylor, W.G.; Cech, A.C.;
 Proc. Natl. Acad. Sci. U.S.A. 88, 415-419, 1991
 A:Title: A broad-spectrum human lung fibroblast-derived mitogen is a variant of hepat
 A:Reference number: A39006; MUID:91110540
 A:Accession: A39006
 A:Molecule type: RNA
 A:Residues: 1-161, 167-728 <RUB>
 A:Cross-references: GB:M55379
 A:Experimental source: embryonic lung
 R:Ioshiyama, Y.; Arakaki, N.; Naka, D.; Takahashi, K.; Hirono, S.; Kondo, J.; Nakayam
 Biochem. Biophys. Res. Commun. 175, 660-667, 1991
 A:Title: Identification of the N-terminal residue of the heavy chain of both native a
 A:Reference number: PH0114; MUID:91207365

A:Accession: PH0114
 A:Molecule type: protein
 A:Residues: 32-43;53-58 <YOS>
 A:Experimental source: Plasma
 R:Weidner, K.M.; Behrens, J.; Vandekerckhove, J.; Birckmeier, W.
 J. Cell Biol. 111, 2097-2108, 1990
 A:Title: Scatter factor: molecular characteristics and effect on the invasiveness of ep1
 A:Reference number: A37796; MUID:91035621
 A:Accession: A37796
 A:Molecule type: protein
 A:Residues: 86-91;329-344;356-363, 'XX', 366-370;425-434;442-447, 'X', 449-450;543-546, 'X', 5
 R:Nakamura, T.; Nishizawa, T.; Hagiya, M.; Seki, T.; Shimomishi, M.; Sugimura, A.; Tashir
 Nature 342, 440-443, 1989
 A:Title: Molecular cloning and expression of human hepatocyte growth factor.
 A:Reference number: S06794; MUID:90066676
 A:Accession: S06794
 A:Molecule type: mRNA
 A:Residues: 1-31, 'HR', '34-77, 'N', '79-292, 'V', 294-299, 'M', 301-316, 'A', 318-335, 'K', 337-386, '
 A:Cross-references: EMBL:X16323; NID:g32081; PIDN:CAA34387.1; PID:g32082
 A:Experimental source: liver
 A:Note: the authors translated the codon CAG for residue 727 as Glu
 R:Hartmann, G.; Naldini, L.; Weidner, K.M.; Sachs, M.; Vigna, E.; Comoglio, P.M.; Birchm
 Proc. Natl. Acad. Sci. U.S.A. 89, 11574-11578, 1992
 A:Title: A functional domain in the heavy chain of scatter factor/hepatocyte growth fact
 A:Accession: I59214; MUID:93087571
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-288, 'ET', 'CHAB>
 A:Cross-references: GB:I02931; NID:9184033; PIDN:AA52649.1; PID:g184034
 R:Myazawa, K.; Kitamura, A.; Naka, D.; Kitamura, N.
 Eur. J. Biochem. 197, 15-22, 1991
 A:Title: An alternatively processed mRNA generated from human hepatocyte growth factor g
 A:Reference number: S15443; MUID:91200041
 A:Accession: S15443
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-288, 'ET', 'CHAB>
 A:Cross-references: EMBL:X57574; NID:g32083; PIDN:CAA0802.1; PID:g32084
 R:Shima, N.; Nagao, M.; Ogaki, F.; Tsuda, E.; Murakami, A.; Hiyashio, K.
 Biochem. Biophys. Res. Commun. 180, 1151-1158, 1991
 A:Title: Tumor cytotoxic factor/hepatocyte growth factor from human fibroblasts: cloning
 A:Reference number: I52253; MUID:92052058
 A:Accession: I52253
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 161-166 <SH>
 A:Cross-references: GB:S62561; NID:g237996; PIDN:AA820169.1; PID:g237997
 C:Genetics:
 A:Gene: GDB:HGF
 A:Cross-references: GDB:I27524; OMIM:142409
 A:Map position: 7q21.1-7q21.1
 A:Insertions: 30/1; 85/2; 123/1; 161/2; 209/1; 249/2; 289/1; 347/2; 390/1; 424/2; 469/1; 48
 C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C:Function:
 A:Description: stimulates mitosis of hepatocytes and other cells
 A:Note: does not have proteinase activity
 C:Superfamily: hepatocyte growth factor; kringe homology; trypsin homology
 C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringe; pyr
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-494,495-728/Product: hepatocyte growth factor #status experimental <ACH>
 F:32-494/Domain: alpha chain #status experimental <ACH>
 F:128-206/Domain: kringe homology <KR1>
 F:211-288/Domain: kringe homology <KR2>
 F:305-383/Domain: kringe homology <KR3>
 F:391-469/Domain: kringe homology <KR4>
 F:495-728/Domain: beta chain #status experimental <BCH>
 F:495-716/Domain: trypsin homology <TRY>
 F:32/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experim
 F:294,402,566,653/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:487-604/Disulfide bonds: #status predicted

Query Match 73.9%; Score 34; DB 1; Length 728;
 Best Local Similarity 55.6%; Pred. No. 62;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VRIGHLYL 9
 Db 629 LRVAAHLYIM 637

RESULT 10

A35644
 hepatocyte growth factor precursor - rat
 N:Alternate names: heparinectin A; scatter factor
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 21-Jul-2000
 C:Accession: A35644; S13211
 R:Tashiro, K.; Hagiya, M.; Nishizawa, T.; Seki, T.; Shimomishi, M.; Shimizu, S.; Naka
 Proc. Natl. Acad. Sci. U.S.A. 87, 3200-3204, 1990
 A:Title: Deduced primary structure of rat hepatocyte growth factor and expression of
 A:Reference number: A35644; MUID:90222197
 A:Accession: A35644
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-728 <TAS>
 A:Cross-references: GB:D90102; GB:M32987; NID:g220766; PIDN:BA14133.1; PID:g220767
 A:Note: the authors translated the codon GAG for residue 70 as Gln, GAC for residue 4
 R:Okajima, A.; Myazawa, K.; Kitamura, N.
 Eur. J. Biochem. 193, 375-381, 1990
 A:Title: Primary structure of rat hepatocyte growth factor and induction of its mRNA
 A:Reference number: S13211; MUID:91031482
 A:Accession: S13211
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-728 <OKA>
 A:Cross-references: EMBL:X54400; NID:g56353; PIDN:CAA8266.1; PID:g4539554
 C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor
 C:Function:
 A:Description: stimulates mitosis of hepatocytes and other cells
 A:Note: does not have proteinase activity
 C:Superfamily: hepatocyte growth factor; kringe homology; trypsin homology
 C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringe;
 F:1-32/Domain: signal sequence #status predicted <SIG>
 F:56-495/Product: hepatocyte growth factor #status predicted <MAT>
 F:56-495/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>
 F:129-207/Domain: kringe homology <KR1>
 F:212-289/Domain: kringe homology <KR2>
 F:306-384/Domain: kringe homology <KR3>
 F:392-470/Domain: kringe homology <KR4>
 F:496-728/Domain: hepatocyte growth factor beta chain #status predicted <BCH>
 F:496-719/Domain: trypsin homology <TRY>
 F:33/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
 F:235,403,569,656/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:488-607/Disulfide bonds: #status predicted

Query Match 73.9%; Score 34; DB 1; Length 728;
 Best Local Similarity 55.6%; Pred. No. 62;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VRIGHLYL 9
 Db 632 LRVAAHLYIM 640

RESULT 11

A60185
 hepatocyte growth factor precursor - mouse
 N:Alternate names: heparinectin A; scatter factor
 C:Species: Mus musculus (house mouse)
 C>Date: 03-Mar-1993 #sequence_revision 26-May-1994 #text_change 16-Jun-2000
 C:Accession: J02117; P02064; A60185; S43416; S45521; S17173; S10966; I48758; J00231
 R:Sasaki, M.; Nishio, M.; Sasaki, T.; Enami, J.

OY 2 RIGHLYT 9
 |||:|:|
 DB 544 RIGHLYT 551

RESULT 14

G69498
 Calcium-binding protein homolog - Archaeoglobus fulgidus
 C.Species: Archaeoglobus fulgidus
 C.Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C.Accession: G69498
 R.Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390: 364-370, 1997
 A.Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A.Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
 A.Reference number: A69250; MID:98049343
 A.Accession: G69498
 A.Status: preliminary; nucleic acid sequence not shown; translation not shown
 A.Molecule type: DNA
 A.Residues: 1-313 <RES>
 A.Cross-references: GB:AE000965; GB:AE000782; NID:g2689288; PIDN:AA889260.1; PID:g264854

Query Match 71.7%; Score 33; DB 2; Length 313;
 Best Local Similarity 62.5%; Pred. No. 39;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VRIGHLYT 8
 : |||:|
 DB 272 IEIGHLYT 279

RESULT 15

JC2361
 melanoma antigen MAGE-3 - human
 N.Alternate names: MAGE 3 protein
 C.Species: Homo sapiens (man)
 C.Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 18-Feb-2000
 C.Accession: JC2361; PH1296; I38438
 R.Ding, M.; Beck, R.J.; Keller, C.J.; Fenton, R.G.
 Biochem. Biophys. Res. Commun. 202, 549-555, 1994
 A.Title: Cloning and analysis of MAGE-1-related genes.
 A.Reference number: JC2358; MID:94311935
 A.Accession: JC2361
 A.Molecule type: mRNA
 A.Residues: 1-314 <DIN>
 A.Experimental source: melanoma cell line DM150
 R.Traversari, C.; van der Bruggen, P.; Luescher, I.F.; Lurquin, C.; Chomez, P.; Van Pel, J.
 Exp. Med. 176, 1453-1457, 1992
 A.Title: A nonapeptide encoded by human gene MAGE-1 is recognized on HLA-A1 by cytolytic
 A.Reference number: PH1294; MID:93018875
 A.Accession: PH1296
 A.Molecule type: DNA
 A.Residues: 168-176 <TRA>
 R.Gaugler, B.; Van den Eynde, B.; van der Bruggen, P.; Romero, P.; Gafioro, J.J.; De Pla
 J. Exp. Med. 179, 921-930, 1994
 A.Title: Human gene MAGE-3 codes for an antigen recognized on a melanoma by autologous C
 A.Reference number: I38438; MID:94157413
 A.Accession: I38438
 A.Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 1-314 <RES>
 A.Cross-references: EMBL:U03735; NID:g468825; PIDN:AA17446.1; PID:g468826
 C.Genetics:
 A.Gene: MAGE-3
 C.Superfamily: tumor associated protein MAGE
 F*168-176/Region: HLA-A1 binding #status Predicted

Query Match 71.7%; Score 33; DB 2; Length 314;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 IGHLYT 8
 |||||
 DB 172 IGHLYT 177

Search completed: June 20, 2001, 14:04:33
 Job time: 80 sec